White Paper Report

Report ID: 98540

Application Number: HD5042008

Project Director: Matthew Gibson (mgibson@virginia.edu)

Institution: American Association for State and Local History

Reporting Period: 9/1/2008-9/30/2011

Report Due: 12/31/2011

Date Submitted: 12/16/2011

Internet Digital Encyclopedia Alliance:

Toward a Community of Practice

Initial Findings on Best Practices for Digital Encyclopedias

Doug Barnett, Bob Beatty, Matthew Gibson,
Uri Nodelman, Ann Toplovich, Jamil Zainaldin

8/26/2009 First Draft
2/9/2011 Second Draft
5/27/2011 Third Draft

Table of Contents

Background

Over the past decade, a number of encyclopedias focusing on geographic areas of the United States have come online. The early successes of projects such as the online version of the *Encyclopedia of Cleveland History* (1998) and the *Handbook of Texas Online* (1999) have been followed by other ambitious initiatives, including the *Tennessee Encyclopedia* (2002), *HistoryLink: The Free Online Encyclopedia of Washington State History* (2003), *The New Georgia Encyclopedia* (2004), and the online version of the *Encyclopedia of Chicago* (2005). These pioneering projects have successfully demonstrated that Internet-based encyclopedias offer significant opportunities for enhancing public access to, and understanding of, humanities content. They have engaged large numbers of users, expanding the audience for this important body of information and increasing the level of public participation. They have demonstrated that authoritative encyclopedic information can be converted to and created for a digital medium, that content of rigorous quality can be solicited and prepared for publication on the web, that readers will use the online material (often in preference to print), that this method of publishing facilitates greater interaction between publisher and audiences with consequent improvements in the ongoing revision of the product, and that the audience for this information is substantial and transcends geographically localized audiences.

Many other geographically based projects—city, state, and even regional in focus—are underway or in the planning stages. For a variety of reasons, however, relatively few projects have managed to go "live" on the Internet over the past decade, despite the substantial encouragement provided by the NEH and other sponsors. This situation reflects in part the fact that encyclopedia projects—whether print or digital—are daunting undertakings requiring long-term commitments of capital and intellectual resources. The long gestation period of these projects is also due to the need to develop workflow practices, technological standards, and distribution infrastructure—components of the production process which could be standardized across comparable projects. That these practices and standards are developed independently points to an inefficient utilization of resources, particularly as sponsoring agencies fund the reinvention of technologies and processes that have already been developed by other projects. Similarly, the lack of commonly agreed-upon technological standards precludes the possibility of interoperability among different projects, limiting their collective effectiveness.

IDEA was formed at the 2006 AASLH annual meeting by participants in encyclopedia projects that are primarily geographical and historical in focus within the United States and its territories. From that meeting and subsequent dialogue, including substantial anecdotal evidence of project needs, IDEA articulated a digital start-up grant proposal to develop a network of encyclopedia producers and those interested in the production of online encyclopedias in order to address these issues. With support from the NEH, the project was launched in September 2008 with the creation of a project team including the principal collaborators, project staff, and a board of advisors (*see Appendix 1*). An online survey of project parameters and practices was developed in the fall of 2008 and communicated to thirty-five projects, with 20 projects responding (*see Appendix 2*). Data from the survey provided the framework for a two-day summit held in March 2009, during which representatives from 13 invited projects reviewed the survey findings and discussed the needs of the community. A second summit with 11 participants was held in January 2011 to review the draft white paper and discuss the community's ongoing needs. From the survey data, the summit discussions, and subsequent discussions with the board of advisors, the principal collaborators have formulated this white paper. General observations have been augmented with case studies from selected projects (*see Appendix 5)*, and we are especially grateful for the contributions from Becky Calwell (West Virginia), Kelly P. Caudle (Georgia), and Brett Thompson (Connecticut).

Digital Encyclopedias:
Purpose and Principles

Information is critical in the 21st century. It is also abundantly available in digital format and through multiple types of media, including of course the printed word. Though ubiquitous and at times presenting itself as authoritative, information can claim nothing more for itself than its name implies. The fact remains, however, that information on the Web can be and often is presented—and received by its audiences—as trustworthy or authoritative when it may not be (or worse). At least as far as authenticity is concerned, the capacity of audiences to make judgments becomes supremely important, as does the availability of other sources on the Web that users can turn to as indisputably reliable. In the environment of the Web, reliability is becoming increasingly important.

The critical distinction between *information* on the Web and the content of an encyclopedia is *context* and *knowledge*, which is the encyclopedia's domain. Like its printed counterpart, digital encyclopedia content is based on a methodology and expertise that bring "added value" to information. In its coverage of subject matter, an encyclopedia strives to be comprehensive. A digital encyclopedia entry is based on a command of relevant facts, a critical understanding of information sources and bibliography, an informed appreciation of what is significant (and why) within a given subject area or topic, and accountability (the author and sources are identified). Professional editorial oversight ensures that all content is reviewed, edited, and checked for accuracy, and that updates or corrections are made to entries already published to the Web. Finally, knowledge-based content means that informed interpretations can be made; this is an important function of an encyclopedia that is distinguishable from bias or personal opinion or even no opinion at all.

For all these reasons, it is critical that digital encyclopedias continue the strong commitment to quality prose that has long characterized their print forbears. It is a commitment that goes beyond consistent copy-editing and word usage to focus on establishing a coherent editorial voice and perspective across the entire project. This insures that readers encounter a reasonably uniform approach to description and explication across categories and levels of entries, that the structure and length of entries are appropriate to the subject, and a disciplined balance between authorial voice and editorial style and policy.

An interoperable, online network of digital encyclopedias developed in accordance with a common content standard grounded in excellence is a tremendously powerful vehicle for legitimating and disseminating reliable knowledge-based content on the Web. Such a network can also play a pace-setting role on the Web, educating users and encouraging other producers of content to aspire toward (and expect) excellence on the Web.

Core Concepts

In current practice, the term 'Digital Encyclopedia' refers to systems that are heterogeneous in scope and functionality. These systems might include digital media, text, and metadata storage systems, reference-linking systems, and content management systems. Thus the term 'Digital Encyclopedia' has been overloaded to the point where it means little perhaps because:

- as yet there is no agreement on what a Digital Encyclopedia is and what functionality is associated with it and/or
- there has been no formal discussion or understanding of them as heterogeneous systems composed of multiple moving parts

Given issues of sustainability and the similarity of project needs across the digital encyclopedia spectrum, thinking and imagining how we might pool resources, ideas, tools, and lessons learned is crucial to the success of these decentralized projects. The following are some of the more logical areas for potential sharing and collaboration:

- Technology reuse: whether using/reusing whole platforms, applications, or tools, this is certainly an area where costs can be cut and efforts toward interoperability be bettered. Again, for certain pieces of technical collaboration and tool sharing to work, the projects will need to embrace some method of standardization whether it be through a common suite of functionalities desired, overall workflow, basic metadata elements, etc.
- Standards creation: based on the knowledge shared across digital encyclopedia projects, conceiving of and communicating about standards (something this whitepaper is perhaps a first step toward) for content description, markup, workflow, etc. would be beneficial for any number of reasons, including cost reduction, content interoperability, and sharing of technology.

This whitepaper attempts to provide a reference model that circumscribes and consolidates what the community's disparate notions of 'Digital Encyclopedia' might be. In the effort to accomplish this, the IDEA participants who attended the summits in Nashville (*see Appendix 3*) identified core concepts that digital reference works need to understand, consider, and realize as they undertake planning and conceptualizing a digital encyclopedia.

1. Goals & Context

As an organization begins to think about creating an online encyclopedia (or any type of digital project for that matter), answering questions such as "why do we want to do this?"; "who needs this?"; and "do we have the funds and access to the infrastructure to accommodate this endeavor?" are important to consider and answer before one begins to drift out into the sea of digital content creation, curation, delivery, and maintenance. Outlined below are further investigations into these challenging questions that also identify questions about the policies, procedures, and questions of sustainability that an organization must likewise consider as it develops what amounts to the "brand" for the online project.

A) Purpose, Values, & Mission

One can think of "purpose, values, and mission" by considering a set of questions that, when answered, should inform decisions about all the other aspects of the encyclopedia; in effect, the answers to these questions become the project's brand—the identity one conveys to the public. It is important for an encyclopedia to have an articulated purpose, a set of values by which it is constructed and delivered, and an overall mission in order to help determine an identity by which questions related to policy, content, and infrastructure can be answered. Some questions related to identifying and articulating the "purpose" of a digital (or even non-digital) encyclopedia are:

- Why is this encyclopedia being created?
- Who is this for—who are the readers, users, and evaluators of the work?
- What are the criteria by which the encyclopedia will be deemed successful?
- What are the scale and scope of the envisioned project?
- How do we evaluate the success of these projects and the content particularly in regards to the stated purpose and audience of these projects?

B) Institutional Partners & Other Relationships

There are a number of reasons and purposes for establishing partnerships or other types of formal and informal relationships with other institutions and organizations. Partnerships can provide an array of benefits ranging from technical to publicity, marketing, and content opportunities. When considering what types of partnerships might be beneficial, it is important to understand the current landscape of relationships that your organization currently maintains. Combining forces with other agencies and organizations which have similar missions can benefit content creation and reuse; leverage possible donor/ grant opportunities (showing that a project has buyin from other organizations provides an imprimatur that can be potentially beneficial); and technological maintenance and sustainability.

If a project is going to partner with other organizations for content reuse, clear understandings of expectations/boundaries, etc. must be established (e.g. if a project were to partner with an image archive to reuse visual material, how might a MOU be constructed that is mutually beneficial and cost effective? If a library became a partner to help preserve archived periodic editions of the project, what would the expectations from both parties be and, again, what would be the mutual benefits? What are the term limits of the agreement? Who are the responsible point people? What are the responsibilities themselves?)

In the realm of technology, establishing relationships with other institutions that have a more robust technical infrastructure and capacity (e.g. colleges/universities) can provide assurance and security as far as server hosting and maintenance. This will rarely be a "cost-free" partnership. There will be recognizable and real storage, backup, and maintenance costs and the project will become subject to certain policy decisions that it does not necessarily control. This is why establishing the terms of the relationship in writing and memoranda is essential.

C) Staff, Users, & Other Stakeholders

Staff

Considerations of staff makeup are in large part dependant upon balancing the mission and goals of a project with the financial viability and strength (both short and longterm) of the organization that is creating the digital encyclopedia. To help determine staffing, a project should consider the following:

- What are the tasks that must be performed to deliver successfully the product as outlined under the purpose and mission statement?
- How many people are needed to perform these tasks efficiently and effectively?
- Are these positions full-time, part-time, contracted? What are the drawbacks and benefits of having a full-time staff (e.g. project investment) vs. contractual work-for-hires?
- Will certain positions be phased out over time based on needs; will new positions be created based on needs?
- Are contributors volunteers, is this part of their job, or are they paid separately?
- How will technology maintained, grown, and enhanced over time? Will there be staff responsible for upgrading the code base for content management and delivery?
- Likewise, how will the design of the site be maintained and refreshed over time? Will this be contracted out on an as-needed basis?

Users

There are many types of encyclopedia users. There are back-end users (content creators, system administrators, application developers and programmers and designers), and front-end users (the audience consuming the encyclopedia's published information—these can be real people or machines/algorithms/crawlers that are doing something with the published data). For discussion here, "user" refers to the public human end-user.

The question of who the targeted human end-user of your resource is ties back to the purpose that the particular encyclopedia is trying to address and the "brand" the encyclopedia is attempting to convey. While one might say that the "general public" is the core audience of the resource, in general, projects will be more successful with a particular constituency in mind. So in thinking about users, here are some useful questions:

- Who is the primary audience of your resource? Teachers? Students? Scholars?
- How does one determine what this representative constituency needs, wants, and will use: focus groups? forums? online feedback?
- How does a project ensure that its design is intuitive and fresh for users: usability testing?
- What kinds of expectations exist on the part of users? (e.g., how long should/will users wait for downloads? how many clicks must they go through to find content?)
- Will users be able to contact encyclopedia staff to report errors or technical problems, offer comments, suggest new content?
- Who will monitor such communication and respond to users?
- Will the resource be open to public discussion blog style, collaborative editing/content contribution, or otherwise?

Other Stakeholders

There are a number of potential stakeholders with whom relationships must be maintained at a variety of levels. These include content contributors, board members (editorial, organizational, technical, etc.), state/federal government officials, and donors (individual, foundation, or corporation). Contact with these stakeholders cannot be overemphasized as many projects will be utilizing these donor bases to sustain the project's existence.

2. Administration & Management

A project's administration defines the tasks, policies, and processes that create and implement the encyclopedia through governance (strategic planning, staffing decisions, etc.), funding and budgetary decisions, internal and external communication (marketing) strategies, and legal (rights and licensing) decisions.

A) Governance

Digital encyclopedias, like all projects, require rules and conditions to run properly and be managed efficiently and effectively. The following are questions and/or issues that should be considered:

- Who will be the person who has the overarching vision and plan for the project and, therefore, be accountable and responsible for its successes and failures? [*Note: this suggests an individual would be responsible which, for the sake of best practices, makes most sense to the creators of this document; however, there are certainly different models of governance that could be considered depending upon the culture, priorities, and needs of the organization: collaborative, top-down, bottom-up, etc. However, the buck typically has to stop at some one individual for purposes of a project's effectiveness and for the sake of accountability.*]
- Who will lead strategic planning initiatives and what will the intent of strategic planning be for these projects?
- What are the rules and procedures that govern policy development? (Meta-Policy?)

- How will the project be staffed (see "Staff" above)?
- How will the success of the project (staff, content, technology, etc.) be defined quantitatively and qualitatively and how will that success be evaluated?
- Who will handle administrative and fiduciary responsibilities including but not necessarily limited to: hiring and evaluating staff, the management of daily operations, establishing and enforcing codes of conduct, budget planning and oversight
- Who will be responsible for insuring that the project adheres to institutional compliances that effect the running and management of the overall organization (e.g. Sarbanes/Oxley, etc.)
- Since an encyclopedia is, by nature, a long-term project, it is almost certain that unforeseen situations will arise that require some flexibility and changes in decisions about policy, functionality, etc. How might that flexibility be integrated into the decision-making process?

B) Funding & Budget

Because these projects should never end—or as some mistakenly say, "go into 'maintenance mode'" (as content and technology continually require maintenance, updates, rewriting, and review)—a huge concern that any project should consider is the overall challenge of cost and sustainability. Some considerations and questions include:

- What are the financial impacts (how much will this cost; conversely, how much might it bring in)?
  - How much are initial costs in getting started? (e.g. buying and setting up servers, hiring and paying benefits for staff, selecting and modifying software, engaging editors and contributors, aggregating media).
  - Beyond startup, what are projected annual costs each year for ten years, twenty years, etc?
  - What effective models exist (or do not yet exist) that might bear the costs and time for content upgrading, review, and changes for the future?
  - What will the impacts of those costs be on the lead organization's other priorities/programs (assuming there are others e.g. within a state humanities council or a university department)?
  - In opposition to the cost, what might be the benefits of this project in garnering new/untapped support from different donor bases and expanding access to the larger organization's content/mission?
- What methods will be utilized to ensure the financial, technical, and intellectual (e.g. ensuring that content is up-to-date) sustainability of the project; who will be charged with creating and managing the policies that lead to sustainability?
- What will the business model for the project look like (i.e. how will it generate money, support, and revenue for sustainability)? Questions to consider that have impacts on this larger challenge are:
  - What type of organizational structure do you exist in?
  - What are the potential partnerships/collaborations available for content and technical support?
  - What access to donor bases would the project have/not have based upon the rules and regulations within that organizational structure?
  - What are the benefits/deficits of being a fully independent 501c3, a department of a university, or a program of some larger organization?
- What are ways to generate revenue for the continued editing and creation of these materials particularly in light of the principle that IDEA participants endorse which is that digital encyclopedias maintain free access as WWW information resources?
  - Ads? (do we recommend that we not use ads? What are potential impacts—positive and negative—of using advertisements?)
  - External links to revenue generating resources (e.g. to "buy this book" at Amazon)
  - Value-added repackaged digital content formats and modes of access (e.g. print-on-

demand, eBooks for devices such as the Kindle, course packets, etc.)
  ○ Social web revenue generating applications current? Or future?
• What are sources of funding available to a project that might provide a continual revenue stream to support these activities? (e.g. repurposed, value-added encyclopedia content in other formats or environments; ads; endowment; etc?)

C) Communication

• Who in the organization is responsible for handling communication with media and other publicity matters?
• How often and in what format should communication take place?
• How might social media be utilized and who should take control of this form of communication?
• How will public communications be coordinated with staff, management, and stakeholder communications?

D) Rights & Licensing

Rights and licensing questions must be considered in everything from the content that one publishes (who owns the rights to this and in what format: the author? publisher? no one? And what does the contract look like that establishes those rights?) to the code base utilized to develop, manage, and deliver that content (is it open source? If not, does the developer own the code base (not a great idea) or does the ownership convert to the purchaser of that code base—i.e. the project)? Other questions include:

• What are the legal issues one must consider when making certain decisions about intellectual property, content creation, technology implementations, etc.?
• Under what policy(ies) is the project's intellectual property regulated: OpenContent? Copyright? Copyleft? Or some other Creative Commons license (see http://creativecommons.org/)? And how do such policies effect partnership agreements and business model possibilities? What are the benefits/deficits of each? How does one treat contributors of content in each model: works for hire? Do contributors retain ownership of their work?, etc.
• How will contracts be managed to secure such things as information objects (text, media, etc.) and what will those contracts stipulate in regards to intellectual property, deadlines, etc.

3. Content

Digital encyclopedias are content vehicles. Consequently, there are a multitude of decisions to discuss and consider related to the methods and workflows in constructing, maintaining, and preserving content.

A) Content Development

An encyclopedia's information objects taken in isolation convey just that, information; however, taken in aggregate with other related content objects these collections of objects do convey context, meaning, interpretive value, and knowledge. The information perceived and potentially interacted with by end-users typically falls under the following categories:

Features and Requirements
• Text and media: the words users read, videos they watch, images they see, maps they explore and plot, the audio clips they hear
• Links: the connective tissue that takes users to other internal/external resources for

further context and information. Like all aspects of the encyclopedia process, editorial decisions and policies need to be made about links:
- ○ when are internal links appropriate?
- ○ which external sites are reliable, appropriate, and "stable"?
- Categories of topics: under what taxonomy is content organized for purposes of content creation, navigation, and refined searching?
- Content inclusion/exclusion criteria: what content is included/excluded based on your purpose and scope? (Perhaps it is more useful to think from the perspective of what a project should exclude and identifying a rationale for that exclusion)
- Intellectual Integrity: how is "intellectual integrity" defined for the particular project?
  - ○ Editorial Governance: how is the process of intellectual integrity generally evaluated and how does one ensure that the process works?
  - ○ Peer-review: who are the peers doing the reviewing and what are the substantive tasks with which they are charged?
  - ○ Fact-checked: what resources are used: primary and/or secondary? Are multiple sources checked for verification?
  - ○ Copy edited: which style is used for consistent presentation of information?
  - ○ Plagiarism: how are entries tested/checked for inappropriate reuse of content?
  - ○ Links: ensuring that external (and internal) content is up-to-date and dependable
  - ○ Change Management:
    - ▫ How will a project handle typos, errors of fact, interpretation, and overall revision/updates?
    - ▫ Who can report errors and by what means?
    - ▫ How are errors and changes verified, implemented?
    - ▫ Will users know when content has been altered (any change or certain kinds of changes?) since the original publication of the content?
- Will entries be maintained and kept up-to-date or written once and "finished"?
- Is this a repurposing of a pre-existing print encyclopedia and if so, what other added values/needs/functionalities does one need to consider when moving into the digital realm?
- For digital encyclopedias, will the project be strictly online or will there be a print version (e.g. print-on-demand) and/or other repurposed dissemination formats (digital or not)?
- How much technical support is available and who has access to it (i.e., do editors? authors? end-users?)
- What portion of the content must be complete and web-ready for the project to launch?

Workflow

One can think of workflow in terms of a charted life-cycle of each article/entry/information object that is part of the encyclopedia. At each point in the process there should be a clear understanding of who needs to do the work at that point (task), how long they have to do it (deadline), and what the next step is (movement). Each encyclopedia will, of course, have its own sets of specialized parts of the workflow (differences in this might be most apparent between projects that are "digitally-born" and projects that are digitized from existed print resources), but there are some common features that we find many to share.

Each step in the workflow must have an actor and someone to whom that actor is accountable. Questions related to workflow and content life cycle that one needs to consider are:
- Who is responsible for or allowed to do the work of a particular step in the workflow? what individual, group, or groups?
- To whom is the responsible person accountable? to whom do they give their work when they are done?

Here are the steps in a fairly general workflow:
- Suggestion: a topic is suggested as one that should start the article life-cycle

- ○ Ideally, topics are suggested with reference to significance/topical coverage guidelines and processes established and disseminated by the project.
- ○ Suggested topics are evaluated by an editorial entity (staff editor, section editor, advisory editor, etc.) according to the project's significance/topical coverage guidelines.
- ○ Suggested topics are accepted, rejected, or flagged for future consideration; reasons for decision are logged; and suggestors are notified of decisions.
- Inception: the topic is accepted to start the article life-cycle
  - ○ Initial decisions about scope, length, due date, appropriate area of editorial insight are made and article enters the project's workflow.
- Invitation/Application: authors are suggested (or suggest themselves) for the article
  - ○ Ideally, authors are identified or suggest themselves with reference to an author selection guidelines established and disseminated by the project.
  - ○ A designated editorial entity communicates with prospective authors and determines whether potential authors meet project requirements and are willing to work within project expectations.
- Commission: an author is attached to the article
  - ○ Author is provided with relevant assignment information (topic, due date, length, guidelines, etc.) and provides confirmation through contractual agreement.
- Creation/Update: the author creates a draft of the article (or a draft of an update to an existing article)
  - ○ Author works within topical and technical guidelines provided by project.
- Submission: the author submits the draft
  - ○ Along with any ancillary information as called for by article guidelines (keywords, illustrations, references, etc.)
  - ○ Per submission procedures established and disseminated by the project.
- Evaluation: the draft is evaluated according to the standards of the encyclopedia. Procedures will vary according to the needs of individual projects but may include any or all of the following. Generally, these are conducted sequentially, with appropriate rounds of revision and resubmission; some projects might switch the sequence of fact checking and copy editing depending upon approach:
  - ○ Initial review for adherence to assignment and overall coverage of topic, usually conducted by a project editor, section editor, etc.
  - ○ Fact check to confirm factual accuracy, verify internal and external links, and address any specific questions raised in the initial review
  - ○ Copy edit to project standards for correctness, clarity, stylistic consistency, authorial voice, language levels, etc.
- Revision / Resubmission / Re-evaluation until draft is acceptable
- Final Preparation: the article undergoes any modification necessary for publication purposes
  - ○ A pre-production process preparing the article for the specific output mechanisms intended for the project.
- Final Examination/Review: the final preparations are evaluated to confirm everything is ready to go
  - ○ Per procedures established and disseminated by the project. It is helpful to have specific checklists to guide and streamline this process.
- Publication: the article is made available to end-users
- Change Management: how do you fit content refreshing/maintenance of entries in this workflow? Some possibilities include:
- User generated changes (corrections, expansions, updates suggested by people using the product, ideally through a process mechanism implemented by the project).
  - ○ Author generated changes identified by the author subsequent to publication.
  - ○ Project driven changes generated by systematic review and evaluation of selected sections of the project.

B) Quality & Accessibility

The criteria and methods used for evaluating content and services. The following are just some of the issues to consider:

- Establishing content's "intellectual integrity" (as listed above: e.g. producing quality content with editorial governance, fact checking, peer review, copy editing, plagiarism checking)
- Link-checking
- Reporting/evaluating programming bugs and system complications (Bugzilla, regular review of error logs, etc.)
- Providing technical support for servers, other hardware, and software over time. Is there access to 24/7 service in case of system problems? Who has technical capacity for administering the servers (doing backups, upgrading operating systems and software/middleware dependencies, etc.) ?
- Complying with Section 508 of the Americans with Disabilities Act
- Utilizing tools and mechanisms to gather qualitative feedback on effectiveness of user experience (i.e. usability testing)
- Tracking and monitoring web access (Google Analytics, wustats, etc.)
- Evaluating effectiveness of encyclopedia in connecting readers with content using different tools and features (e.g. web statistics analysis, "rate this article" or other social web tools, etc.)
- Implementing methods for content change management

C) Metadata

When we refer to "metadata" in the context of web-based projects, we typically mean some type of platform-independent method of ASCII markup, a system of elements and attributes to describe content (content here can be other ASCII or binary text, still and moving images, audio files, etc.). The form of markup could be eXtensible Markup Language (XML) or Standard Generalized Markup Language (SGML) of which XML is a stricter subset. For the sake of economy, power, and efficiency, IDEA recommends using some application of XML. Some general questions to consider about metadata are:

- How much metadata should content have? The answer to this really relies on what method you use to extract data about your content, the level of accuracy to which you can relate that information to other data, how much you are hoping to use your content for functions such as searching, browsing, and mapping, and how integrated you want your content to be. IDEA recommends projects weigh the functionalities desired, metadata required to enable those functionalities, and the cost, benefit, and overhead of adding and maintaining that metadata. There are those who argue that metadata is a ridiculous failure and those that argue for extreme markup. In general, IDEA suggests projects take a pragmatic and economic approach to deciding where they are along this continuum.
- What kind of metadata should content have? Again, this largely depends upon what you want to enable in your content. In theory the richer and smarter the metadata, the more you might do. Here are some textual features one might consider describing with markup:
  - Creator of resource
  - Title of resource
  - Dates
  - Places
  - Names of people
  - References to other encyclopedia entries
  - General structures (paragraphs, quotes, etc.)
- Which schema should be used? IDEA does not recommend a particular markup application or combination of applications at this time; however, it is strongly suggested

that projects employ existing markup standards that are developed and maintained by national and international user communities and that whatever markup one chooses that it should maintain a separation of form from the content—i.e. mark up what the content is not how it should look or display on the screen (*see Appendix 4 for a brief discussion about and links to possible and appropriate standards*).

D) Archiving

Based on variables such as cost and benefit, what are acceptable levels for securing and preserving the content and service? Determining best methods for archiving material for future use—both in the case of technological evolution (monitors/tools utilizing higher resolution visual material) as well as content corruption/destruction—must be a part of the planning process for a digital project of this nature. Some questions of archiving and support might include:

- Is there 24/7 availability in the server room in case of problems/interruptions in service? Can we partner with another organization to help us accommodate this need or do we need to hire a firm for this service?
- What type of backup policy do we implement? Daily and incremental backups for preservation of content? Where is the media stored that holds these backups?
- In the case of larger disaster, do we geographically distribute content servers through mirroring and, again, where would the backups need to reside in case of natural disaster?

4. Infrastructure

A) Physical

- Will new functionalities and technologies be added to the site over time as they become available and is the digital encyclopedia system currently in place flexible enough to accommodate extension with little time, effort, and, most importantly, money?

B) Hardware

Hardware includes everything from the client computers and local storage that editors and administrators use to manage and maintain the digital encyclopedia to the actual servers and online storage where the content is kept, managed, and delivered. Considerations related to hardware include:

- Cost: for initial purchase, housing, as well as the ongoing lifecycle of upgrades and replacement costs
- Capacity: how much storage (local and remote) will a project need in the short and long term; will a project need two or more separate machines to run a database server and web server, etc.
- Security and stability: what will a project require in hardware to ensure content redundancy and uninterrupted access
- Speed: what types of network switches and CPU speeds should a project require of its system to ensure speedy response times to user/client requests
- Compatibility/ability to host/operate software needed for project.

C) Software and Publishing Systems

As with most things involving technology, there are always more than one solution to a challenge. The question of "what software or platform do we use for managing and

publishing our content" is less relevant than what your functional requirements are for that content and how your users (editors, administrators, and human/non-human public users) interact with that content. There are, however, some basic policies that one should keep in mind when thinking of a publishing platform for content:

- How much does this software cost (initially and in the future as it needs maintenance and updating) and how much is the project willing to spend?
- Who owns this software and where is the code base housed/maintained? What provisions are in place in the event that the software provider goes out of business? If the code base itself is open source what complications might a project have down the road if the source code is modified for the project and subsequently new versions are released?
- Will a project want to invest time and money into someone on staff to be responsible for software/feature additions and maintenance?
- Ideally a project will want to have access to and control of the source code. A well-documented and open standard code base is incredibly important for future maintenance and extensibility.

D) Support

Besides questions of software and hardware, more concrete and practical questions to consider in relation to technology are:

- What policies and contingencies does a project have in place to ensure content safety and system stability?
- How can one ensure continual access to content even when the primary servers encounter system/network problems or storage fails? What policies of content redundancy does a project maintain (site mirroring; load balancing; etc.)?
- What resources exist (either contractual or institutional) to help in housing the technology and in assisting with technical upgrades (both software and hardware)?
- Depending upon the server hosting arrangement, what guarantees of site and information stability and perseverance does a project have (e.g. in case entity goes out of business, in case of natural disaster, in the case of etc.)?
- How often should a project expect to replace/add on to hardware and upgrade existing servers? (How will the project plan for hardware and operating system/database/web server obsolescence?)
- How will the project plan for the possibility of software upgrades and major and minor enhancements to the site's design, functionality, or the content management? (How will the project plan for design and software obsolescence?)
- What budgeting does a project need to have in place to implement these changes and upgrades?
- What are the project's obligations to comply with Section 508 of the Americans with Disabilities Act? IDEA recommends that all sites adhere to Section 508 regardless of organization's business classification.

5. Conclusion

In the world of print, the task of building and publishing an encyclopedia follows some well worn paths. A digital encyclopedia is another matter, for the subject-area expertise that is the hallmark of any encyclopedia becomes but one of many requirements and considerations.

In what ways can an encyclopedia fulfill its essential purpose by making use of the multitude of opportunities presented in a digital environment—opportunities that seem to change daily? Certainly this was a key question throughout our conversations and discussions, but so also was the question of how the concept of the encyclopedia is being influenced and even changed by the very existence of this environment. Do we include as subjects individuals who are still living? Do we provide external links? Do we offer interactive options for users?

Should every version of an article published to the Web be archived? Should the entire site? Who are the primary audiences of online encyclopedias, and what do users expect of an encyclopedia online? What kind of content management system should be used?

There is also the question of how to sustain an encyclopedia when it is launched and the responsibility of maintaining the site's currency. The human, financial, organizational, and technological investments for developing, hosting, maintaining, and updating a site often are beyond the capacity of any single organization, making new kinds of partnerships and collaborations an important avenue for exploration early in a project's history.

How do we make the best use of unfolding opportunities at the planning stage? How do we govern projects that can take on lives of their own? How do we ensure that the labor of all our efforts—the creation of content and the means for its delivery and use—is sustained? The advantage of traditional publication is that, when done, the volume(s) sits on the shelf. That is also a disadvantage.

There is no one model, we are learning, but rather consequential choices of staffing, technology, methodology, content, funding, and sponsorship that narrow or enlarge other options along a decision-making tree. Ideally, an online encyclopedia is never complete because knowledge and technology do not sit still; nor do audiences.

The task before our colleagues who convened for focused presentations, discussion, and reflection in Nashville in 2009 and again in 2011 was to create a "community of practice" that was known by its shared goals, values, and commitment to online accessibility. We heard from those who had undertaken online encyclopedia projects, and learned from the problems they encountered and the decisions that were made. We also learned from what worked. We heard from technology experts about the future of the Web and emerging tools, and from librarians, scholars, project managers, authors, and editors of online encyclopedias. *Toward a Community of Practice* captures what we learned and sets forth our collective vision of best practices for the next generation of encyclopedia projects.

We wish to express our appreciation to the American Association for State and Local History (AASLH) for their encouragement and support as host of the Internet Digital Encyclopedia Alliance (IDEA). IDEA was created in 2006 at an AASLH national conference. In subsequent years AASLH continued to host workshops and informal discussion groups amongst IDEA's affiliate members, creating a sense of community and shared endeavor. We would like to thank especially Terry Davis, President and CEO of the AASLH, whose enthusiastic support made this project possible. We would also like to thank Bob Beatty, Vice President of Programs for AASLH, for his outstanding project management and leadership. The participants at our two dialogues in Nashville are the trailblazers of online encyclopedias and Web-based educational resources; we are enormously grateful for their time and their willingness to share their wisdom, experience, and lessons learned.

Foremost, the generous support of the National Endowment for the Humanities made our conversations in Nashville possible, offering time for focused presentation and conversation that was essential for the preparation of this report. Among our project goals is the creation of a community of practice defined by agreed upon content and technical standards. *Toward a Community of Practice*, we hope, will become a stepping stone.

Appendix 1: Staff, Advisors, and Collaborators

Project Staff

Beatty, Bob, Director of Programs, American Association for State and Local History, Nashville, TN

Davis, Terry, President and CEO, American Association for State and Local History, Nashville, TN

Gibson, Matthew S., Managing Editor, *Encyclopedia Virginia*, Virginia Foundation for the Humanities, Charlottesville, VA

Board of Advisors

Graham, P. Toby, Director, Digital Library of Georgia, University of Georgia, Athens, GA

Grossman, James, Vice President for Research and Education at the Newberry Library and Senior Research in History at the University of Chicago, *The Encyclopedia of Chicago*, Chicago, IL

Halbert, Martin, Director of Digital Programs and Systems, Emory University Libraries, Atlanta, GA

Nodelman, Uri, Engineering Research Associate at the Center for the Study of Language and Information and Senior Editor of the *Stanford Encyclopedia of Philosophy*, Stanford University, CA

Reiff, Janice, Associate Professor of History, UCLA, and *The Encyclopedia of Chicago*, Los Angeles, CA

Seaman, David, Associate Librarian for Information Management, Dartmouth College Library, Dartmouth College, Hanover, NH

Principal Collaborators

Barnett, Douglas E., University of Texas Libraries and *The Handbook of Texas*, Austin, TX

Gaventa, Matthew, Media Editor, *Encyclopedia Virginia*, Virginia Foundation for the Humanities, Charlottesville, VA

Toplovich, Ann, Executive Director, Tennessee Historical Society and Managing Editor, *The Tennessee Encyclopedia of History and Culture*, Nashville, TN

Zainaldin, Jamil, President, Georgia Humanities Council, and *The New Georgia Encyclopedia*, Atlanta, GA

Appendix 2: Survey of Project Practices and Parameters

Projects Surveyed (2008)

Encyclopedia of Alabama (http://www.encyclopediaofalabama.org/)

Encyclopedia of Arizona (in development)

Encyclopedia of Arkansas History and Culture (http://www.encyclopediaofarkansas.net/)

Encyclopedia of Chicago (http://www.encyclopedia.chicagohistory.org/)

Laptop Encyclopedia of Connecticut History (http://www.ctheritage.com/encyclopedia/encyclopedia.htm)

New Georgia Encyclopedia (http://www.georgiaencyclopedia.org/)

Guampedia (http://www.guampedia.com/)

HistoryLink: The Free Encyclopedia of Washington State History (http://www.historylink.org)

Kentucky Encyclopedia (http://www.kyenc.org/)

KnowLA: Digital Encyclopedia of Louisiana History, Culture, and Community (http://www.knowla.org/)

Mississippi Encyclopedia (http://www.olemiss.edu/depts/south/ms_encyclopedia/)

Online Nevada Encyclopedia (http://www.onlinenevada.org/)

Te Ara: Encyclopedia of New Zealand (http://www.teara.govt.nz/)

Ohio History Central: An Online Encyclopedia of Ohio History (http://www.ohiohistorycentral.org/)

Encyclopedia of Oklahoma History and Culture (http://digital.library.okstate.edu/encyclopedia/)

Oregon Encyclopedia (http://www.oregonencyclopedia.org/)

Enciclopedia de Puerto Rico (http://www.enciclopediapr.org/)

South Carolina Encyclopedia (currently a print encyclopedia looking to digitize)

Standford Encyclopedia of Philosophy (http://plato.stanford.edu/)

Dictionary of Sydney Project (http://www.dictionaryofsydney.org/)

Tennessee Encyclopedia of History and Culture (http://tennesseeencyclopedia.net/)

Texas Handbook Online (http://www.tshaonline.org/handbook/online/)

Encyclopedia Virginia (http://www.encyclopediavirginia.org/)

West Virginia Encyclopedia (http://www.wvencyclopedia.org)

Executive Summary of Survey Results[1]

In general, this questionnaire was meant more to get a sense of the online encyclopedia/reference landscape than anything else and to give all of us a point of comparison in how

---

[1] For access to the full survey results, please see http://www.people.virginia.edu/~msg2d/IDEA

we are doing things with our individual projects. I think the information here can be used as a starting point for the summit in Nashville, but we will be going into much more depth in regards to content production strategies and technical development as it relates to cost, scalability, and potential cross-project integration.

One caveat—perhaps there are others—about this survey: where we see a large aggregation of data, say in the response to the very first question, this is not totally accurate since we had a disproportionate number of responses from single projects; apologies for this skewing of data, the number of respondents per project should have been limited to one per question; however, the results still provide a decent point of departure.

Infrastructure, Planning, and Audience Information

- **Lead Organization Type**: Most of the contributions came from individuals who are creating projects under the leadership of academic and/or cultural agencies such as state humanities councils, university/college departments, historical societies, museums, libraries, etc.; no one identified themselves strictly as a "for profit" agency.
- **Project Partners**: Content/technology development partners tended to be permutations of the same types of organizations in the "lead" category as above with the exception that government agencies play a much larger role as a partner than they do as a lead institution. (I think the question on the survey may have been poorly phrased since I imagine many respondents assumed "partner" here to be a financial supporter).
- **Staff Size of Lead Organization**: Total staff size of lead organizations varied widely especially when the lead org was a university department or government agency. The statistic was pretty much split between 14 respondents saying they worked within a lead org that had between 1 and 10 staff members and 6 respondents who worked within an org between 100 and 1000.
- **Full-time Project Staff Size at Production Peak**: Project staff size at the height of project production is somewhat evenly distributed between 1 and 7 full time positions with the majority clumping up around 6 or 7 FTEs with three outliers: one at 0, at 15, and 19.
- **Part-time Project Staff Size at Production Peak**: Part-time project staff at the peak of production was primarily between 3 and 4 with some outliers at either end of the spectrum.
- **Project Staff Size in Maintenance**: In "maintenance" mode, most respondents said that they keep or plan to keep either 1 or 2 FTEs employed; one respondent said "Maintenance mode requires the same personnel and costs, given our mission," and several other respondents stated that they would keep between 0% to 25% FTE dedicated to the project in "maintenance mode."
- **Publication Launch Date**: 70% of the respondents surveyed have "launched" with the majority having gone online between 2004 and 2008. Four respondents anticipate launching in 2010 and four respondents stated that they launched in the 1990s.
- **Encyclopedia Types**: 85% of those surveyed said that their encyclopedia is geographically based. There were no encyclopedias organized around a particular time period.
- **Target Audiences**: The primary target audiences were "Students of grades K-12" and the "General Public." While no one explicitly indicated Businesses, State/Local Government, or Genealogists as a primary audience, two respondents indicated "all of the above" as their answer.

Budget and Financials

- **Total Project Cost**: Most respondents stated that the **total cost (i.e. cumulative production cost over multiple years)** of their project was between 1 and 4 million. Two respondents on indicated their project cost less than 250K and two indicated that their

project was between 6 and 10 million.

- **Annual Maintenance Costs**: The annual budget for maintenance of the project after the core content has reached a "critical mass" was 250K/year for 7 respondents where 5 respondents indicated that their maintenance costs were between 250K-500K per annum and 11 indicated that their costs were between 0 and 100K per annum.
- **Funding Sources**: Primary sources of funding based on most responses come from four sections:
  - ○ Corporation/Foundation Donors (14 respondents)
  - ○ Federal Government: grants, earmarks, etc. (13 respondents)
  - ○ State Government: grants, earmarks (11 respondents)
  - ○ Individual Donors (10 respondents)
- **Access Cost**: 26 respondents said their project does not and/or will not charge for access, two were unsure, and one stated that they may consider otherwise due to lack or loss of government funding.
- **Revenue Generation**: In regards to supplemental methods of revenue generation, 38 responses were pretty evenly split between the sale of print volumes or ancillary print material, sale of electronically repurposed "value-added" material (e.g. eBooks, etc.), sale of images, and use of advertisements (e.g. Google ads). 12 respondents indicated that they were not sure what methods they would use to generate revenue.
- **Contributor Payments**: Payment to contributors breaks down as follows:
  - ○ Pay a flat rate from 100$/entry to some projects paying a graduated rate based on length of entry which ranges from around 50$ to 300$ (10 respondents)
  - ○ Pay by word from 5 cents/word to 40 cents/word (6 respondents)
  - ○ Don't pay at all (6 respondents)
- **Average Contributor Pay per Entry**: For those projects that do pay contributors, the respondents indicated that the average payment to a contributor is between 25$ and 150$.

Content

- **Source of Content**: 14 respondents said the genesis of their content was/is/will be a mix between digitally-born and digitized print resource material; 12 respondents said all of their material was/is/will be entirely digitally born; one said that all material has been digitized from a print publication.
- **Author Choice**: Ideal author choices are listed in the following preferred order:
  1. Academics/Scholars (25 respondents)
  2. Graduate Students (11 respondents)
  3. Other types of recognized experts (8 respondents)
  4. Professional freelancers (6 respondents)
  5. Anyone who wants to write (4 respondents)
  6. Journalists (3 respondents)
- **Content Development Methods**: were somewhat split with slightly more respondents stating that they develop a large all-encompassing list and work from that (16 responses) while slightly fewer respondents stated that they divide the encyclopedia thamtically and work on "sections" of content over time (12 respondents). Several respondents stated that they create content organically with large consideration of donor wishes and market forces. Two respondents stated that they create content chronologically.
- **Editorial Structure**: Most responses indicate the use of one or more full-time editors and an editorial board. 19 respondents stated that they engage one or more "section" or "guest" editors per topic to help provide content direction, etc.
- **Peer Review**: is a large concern among most respondents with 18 stating that they have single or several topical specialists who develop and oversee content of particular sections and 15 respondents stated that they have an editorial board for oversight. One respondent mentioned that peer review was cut from the budget while two respondents stated that they do not use peer review.

- **Quality Control**: 96% of responses indicated that the project uses internal or external copyediting; 84% employ internal or external fact checking; and 68% check hyperlinks; one respondent mentioned doing occasional spot-checking for plagiarism.
- **Inclusion Criteria—Biographies** breaks out as follows:
  1. Person's career must have been spent in and/or had large impact on the respective state/locality (23 responses)
  2. Must be dead (4 responses)
  3. Must be mentioned elsewhere--and therefore have a link--to some other entry in the encyclopedia (4 responses)
  4. Must be born in state/locality (3 responses)
  5. Still not really sure (2 responses)
  6. Must have died in state/locality (1 response)
- **Inclusion Criteria—Places**: 100% of respondents say their encyclopedia includes entries on cities, towns, and villages with some specific criteria (NB: I must say, this was sort of a poorly phrased question, my apologies)
- **Inclusion Criteria—Changing Geographical Boundaries**: 17 respondents stated that they include topics relevant to historical boundaries within respective time period while 3 respondents stated that they include topics relevant only to contemporary geographical boundaries.
- **Style Usage**: The majority of respondents (16) say their project uses a customized house style; the close runner-up was Chicago with one stating that they use a modified version of Chicago (13 respondents); and another uses a non-U.S. government style guide. No respondent mentioned MLA, APA, or Turabian.
- **Entry Length Policies**: 22 respondents said they have policies while 3 said they do not. For more in-depth commentary on those policies, see the full results.

Technology

- **Who Builds Technical Framework/CMS**: Over 50% of responses indicate that at least their project's technological foundation was built using outsourced corporate IT group. 9 respondents (around 37% of responses) stated that their publishing/content management framework was built in collaboration with acadmic/university IT partner. And 6 respondents (approx 25% of responses) say that they built their tech infrastructure in-house.
- **In-house Technical Familiarity**: Based on the answer above, there is not a large number of projects that have in-house technical knowledge of how their system works or was built. That being said, 14 respondents did say that they had someone on staff who could speak with some reliability on how the content publishing system was created and what technologies it employs. Most others (24 respondents) stated that this type of knowledge is housed with outsourced developers or other institutional partners.
- **Web Hosting**: 8 respondents stated that their projects are hosted by independent IT vendor, another 8 stated that their web server is housed/maintained at an affiliate/partner institution, while 4 respondents stated that their projects are hosted in-house.
- **Operating System**: UNIX/LINUX drives most of the online encyclopedia community (13 respondents stated that they run some form of UNIX or LINUX be it Solaris, Fedora, etc.); 5 respondents stated that their server runs on a Windows platform; and 3 stated that their server runs on MAC OSX.
- **Content Management Methods**: The majority of respondents (approx. 20) stated that they use a Content Management System (CMS) or online database of some sort while 4 respondents create and manage content offline and upload to site.
- **CMS Development**: Of those respondents who stated that they use a CMS--
  - 12 responded that their CMS was built from scratch by hired vendor or institutional partner
  - 3 respondents stated that they use but modify to different extremes an off-the-shelf open-source CMS such as OpenCMS or Drupal

○ 2 respondents stated that they use an off-the-shelf proprietary CMS (e.g. Microsoft CMS)

- **Content Storage Method**: Most responses (12) indicate that encyclopedia content is stored directly in a relational database such as MySQL, PostGres, Oracle, Access, etc.; Close in the running (10) describe storing static HTML/XML files directly on the filesystem; where 5 suggested a combination method most likely of storing complex XML inside relational tables; finally two respondents mention using Amazon's Simple Storage Service which uses its own proprietary relational database system.
- **XML and XML Schema**: For those projects that use XML to store or publish content, 6 repondents state that they use a home-grown schema; 5 respondents use Dublic Core; 4 use XHTML; 3 have no clue; 2 use TEI; 2 use METS, one uses MODS, and no one uses DITA or RDF for content storage.
- **Media Usage**: The types of media projects are using to contextualize/add to entry texts is broken down as follows:
    ○ Still images (26 responses)
    ○ Moving images (23 responses)
    ○ Audio clips (22 responses)
    ○ Maps (22 responses)
    ○ 3-D renderings (3 responses)
    ○ No media (1 response)
- **Delivery and Scripting Technologies and Applications**: In order of responses, the most common delivery, client-side, and webpage formatting technologies are:
    ○ Flash (14 responses)
    ○ Javascript (11 responses)
    ○ CSS (10 responses)
    ○ AJAX (5 responses)
- While the most common middleware/scripting languages are:
    ○ PHP (8 responses)
    ○ JAVA (5 responses)
    ○ ASP (2 responses)
    ○ XSLT (2 responses)
    ○ PERL (1 response)
    ○ Python (1 response)

One respondent indicated that they use XQuery to query their native XML database.

- **Searching Tools**: Most respondents indicated that they us a home-grown or purchased proprietary search tool for their content (10); Site-specific Google Search was next (4 respondents), and finally 2 respondents state that their project uses MySQL or PostGres full-text search.
- **Content Integration Technologies**: For the most part, everyone took the 5th on this one stating that they had no clue what to use or how to use what they didn't know what to use. 4 respondents stated that their content would be published so that it could be ingested by OAI harvesters; 2 respondents stated that they have an API which exposes content for others to mash up with their projects; and 1 respondent stated that they use Z39.50 protocols.

Appendix 3: Summit Participants

First Summit: March 2009

Doug Barnett, *Handbook of Texas Online*, University of Texas at Austin

Bob Beatty, American Association for State and Local History

Ben Bernston, *Encyclopedia of Alabama*, Auburn University

Kent Calder, *Handbook of Texas Online*, Texas State Historical Association

Kelly Caudle, *New Georgia Encyclopedia*, Georgia Humanities Council

Terry Davis, American Association for State and Local History

Matt Gaventa, *Encyclopedia Virginia*, Virginia Foundation for the Humanities

Matthew Gibson, *Encyclopedia Virginia*, Virginia Foundation for the Humanities

Pam Heath, *HistoryLink.org*, Independent Consultant

Jeff Jakeman, *Encyclopedia of Alabama*, Auburn University

Pat Kaetz, *Encyclopedia of Alabama*, Auburn University

Anna Lancaster, *Encyclopedia of Arkansas History and Culture*, Central Arkansas Library System

Sarah Marcus, *Encyclopedia of Chicago History*, Chicago History Museum

Tanya Mendiola, *Guampedia*, Guam Humanities Council

Uri Nodelman, *Stanford Encyclopedia of Philosophy*, Stanford University

Steve Portch, *Handbook of Texas Online*, Texas State Historical Association

Maryann Reissig, *Tennessee Encyclopedia of History and Culture*, University of Tennessee Press

Randal Rust, *Ohio History Central*, R. Squared Communications

Nathania Sawyer, *Encyclopedia of Arkansas History and Culture*, Central Arkansas Library System

Tom Scheinfeldt, Center for History and New Media, George Mason University

Harry Searles, *Ohio History Central*, Ohio History Society

Ann Toplovich, *Tennessee Encyclopedia of History and Culture*, Tennessee Historical Society

Harlan Wallach, *Encyclopedia of Chicago History*, Northwestern University

Jamil Zainaldin, *New Georgia Encyclopedia*, Georgia Humanities Council

Follow-up Summit: January 2011

Doug Barnett, *Handbook of Texas Online*, University of Texas at Austin

Bob Beatty, American Association for State and Local History

Ben Bernston, *Encyclopedia of Alabama*, Auburn University

Sheila Brennan, Center for History and New Media, George Mason University

Kent Calder, *Handbook of Texas Online*, Texas State Historical Association

Kelly Caudle, *New Georgia Encyclopedia*, Georgia Humanities Council

Andrea Ferguson, *KnowLa: Encyclopedia of Louisiana*, Louisiana Endowment for the Humanities

Matthew Gibson, *Encyclopedia Virginia*, Virginia Foundation for the Humanities

Erica Hartmann, *Encyclopedia of Minnesota History*, Minnesota Historical Society Press

Pat Kaetz, *Encyclopedia of Alabama*, Auburn University

Mike Keller, *West Virginia Encyclopedia*, West Virginia Humanities Council

Sarah McKee, *New Georgia Encyclopedia*, Georgia Humanities Council

Joyce Miller, *KnowLa: Encyclopedia of Louisiana*, Louisiana Endowment for the Humanities

Uri Nodelman, *Stanford Encyclopedia of Philosophy*, Stanford University

Ken Sullivan, *West Virginia Encyclopedia*, West Virginia Humanities Council

Brett Thompson, *ECHO, Encyclopedia of Connecticut History*, Connecticut Humanities Council

Jamil Zainaldin, *New Georgia Encyclopedia*, Georgia Humanities Council

Appendix 4: Technology and Standards

A Digital Encyclopedia Reference Model

Multiple understandings of "Digital Encyclopedia" results in a lack of interoperability and reuse of both content and technologies. Here we attempt to address aspects of technology that should be considered when thinking of content management, use, and reuse; system needs and architecture; and content preservation and sustainability. The first step is to think of a Digital Encyclopedia in an abstracted conceptual 3-tier conceptual framework:
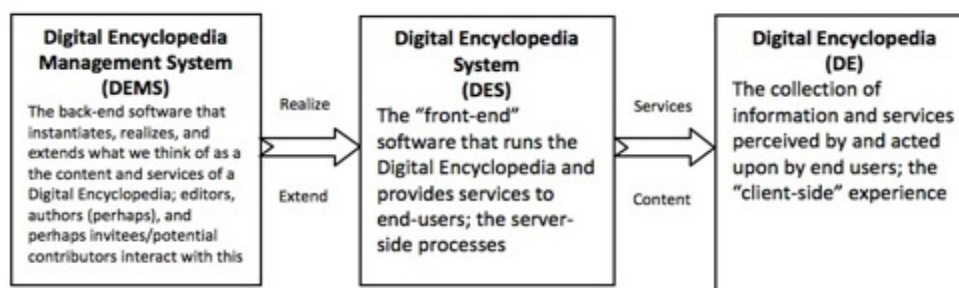
The visual model above suggests breaking down the conceptual framework of the Digital Encyclopedia into three layers which are further described below:

- Digital Encyclopedia Management System (DEMS): the back-end environment that is used to realize and address the needs of end-users and to extend the content and services of a Digital Encyclopedia; end-user here must be understood as the human being (information user) AND the inanimate entity (external computer, software, and other applications—e.g. bots/webcrawlers/API users that interact with the content as data); the DEMS:
    - produces and administers the DES while incorporating the suite of functionality and needs considered fundamental for Digital Encyclopedias (e.g. enables submission, proofing, editing, updating of entries; logs and tracks topics, authors, and deadlines; manage entire project, publish entries/media)
    - integrates additional software that offers more refined, specialized, or advanced functionality
    - this is the layer that editors, potentially authors/invitees utilize to manage content and workflow
- Digital Encyclopedia System (DES): the front-end software that runs the Digital Encyclopedia and provides services to end-users by taking user requests and needs and interacting with the DEMS to return content (e.g. read/play/interact with media/entries/maps, etc.) and/or satisfy the demands the user has (e.g. post/link content as in Blogs, etc.)
- Digital Encyclopedia (DE): a possibly virtual organization of systems that creates and/or collects, manages, and sustains for the long term rich digital content, and offers to its user communities specialized functionality on that content, of measurable quality and according to codified policies

Within these layers, the primary areas of consideration are functionality (system capabilities and services available to different groups of users), architecture (the hardware and software implementation to enable design and functionality), and creation and design (both design and organization of back-end information architecture for content creation, relation, and maintenance as well as the realized front-end pages that interleave functionality within the public interface and interactive experience).

Content Metadata

Again, when we refer to "metadata" in the context of web-based projects, we typically mean some type of platform-independent method of ASCII markup, a system of elements and attributes to describe content (content here can be other ASCII or binary text, still and moving images, audio files, etc.). The form of markup could be eXtensible Markup Language (XML) or Standard Generalized Markup Language (SGML) of which XML is a stricter subset. For the sake of economy, power, and efficiency, IDEA recommends using some application of XML. Some examples of possible XML markup dialects that would pertain to digital encyclopedia projects include:

- XHTML 1.0 (http://www.w3.org/TR/xhtml1/)
- Dublin Core (http://dublincore.org/)
- METS (Metadata Encoding and Transmission Standard http://www.loc.gov/standards/mets/)
- TEI (Text Encoding Initiative: http://www.tei-c.org/)
- MODS (Metadata Object Description Schema: http://www.loc.gov/standards/mods/)
- RDF (Resource Description Framework: http://www.w3.org/TR/rdf-schema/)
- See http://www.digitalpreservation.gov/formats/content/text_preferences.shtml for a general discussion of text creation standards

Media Standards

Considerations of media standards (i.e. images, moving images, audio) should be divided into issues of delivery and challenges of preservation and archiving. The philosophical basis behind rich XML (markup something once so that it is flexible and rich enough to be reused in many current/future applications) is the same philosophy behind the employment of rich media digitization standards. For a synopsis of media digitization practices see:

- Still Image: http://www.digitalpreservation.gov/formats/content/still_preferences.shtml
- Moving Image: http://www.digitalpreservation.gov/formats/content/video_preferences.shtml
- Digital Audio: http://www.digitalpreservation.gov/formats/content/sound_preferences.shtml

For delivery of media objects, one should consider applications that are readily available to targeted audience (e.g. Flash, Quicktime or other client-side media players, etc.)

Content Storage

There are a number of methods that projects might use to store content. In this context, "storage" defines the state in which content is kept that can then be accessed for any number of different procedures and purposes (editing, publishing, reading, etc.) by middleware and/or server-side and client-side scripting. There is a variety of storage management systems and what follows represents options but in no way exhausts possibilities:

- Relational Database: collects information as a set of relations which is to say the information is grouped by common attributes found in the data set. Examples of Relational Database Management Systems include MySQL, PostGreSQL, Oracle, and SQL Server.
- XML Database: allows the data to be stored as XML instead of being sliced up into fields inside of tables. There are two major classes of XML databases that currently exist:
  - XML-enabled: is really a hybrid of database which relies on the relational database for storage and retrieval but in which the data itself is accepted and stored as XML or is mapped to relational fields easily rendered as XML on output. Examples include MySQL, Oracle, PostGreSQL, and SQL Server.
  - Native XML: the internal model of the Native XML database depends on XML and

uses XML documents as the fundamental unit of storage, which are, however, not necessarily stored in the form of text files. Some examples of Native XML databases are eXist, Xindice, and MarkLogic.

- ○ Relational vs. Native XML Databases see Ronald Bourret's page: http://www.rpbourret.com/xml/XMLAndDatabases.htm and for a list of different types of database management systems, see: http://www.rpbourret.com/xml/XMLDatabaseProds.htm
- File System: while not recommended because of processor overhead and the difficulty in managing content updates and versions, one can store content directly on the file system; but, again, this is not recommended; though it does make more sense in the case of media (images, moving images, and audio).

Other Considerations

- Accessibility
  - ○ To learn more about the Americans With Disabilities Act and Section 508 see: http://www.access-board.gov/sec508/guide/1194.22.htm
  - ○ if you decide to employ technologies that require client-side software or plug-in installation (e.g. Flash, Google Earth, AJAX, and other JavaScript requirements) consider some of your core audiences that may not have administrative access to configure and install new software even it that software is free.

Appendix 5: Case Studies

Encyclopedia of Connecticut History Online
A program of the Connecticut Humanities Council

Case Study: Selecting a Technology Partner

A year into the development of the Encyclopedia of Connecticut History Online (ECHO), the Connecticut Humanities Council found itself without a technology partner to continue work on designing and building the website. We had developed a very successful relationship with a group associated with a major insurance company that chiefly provided web consulting services to internal clients but also took on projects outside the organization. Thus, we were fortunate to gain access to top notch talent that is typically beyond the reach of a nonprofit organization. We lost that access, however, when the economic downturn hit and the group was directed to shed its external clients to focus on internal needs.

We had connected with our insurance company partners through word of mouth, and we now found ourselves faced for the first time with a more traditional search. We quickly found that the Internet technology consulting landscape is both vast and varied, making it difficult to find an appropriate resource in a Goldilocks-like world of too big and too small. We began our search, then, by finding websites with a "look and feel" similar to ECHO's and then tracking down the technology consulting firms behind them. This effort produced a list of twenty candidate firms. We then narrowed the field to ten by applying the following set of filtering criteri:

1. Experience with creating information-rich websites;

2. Experience working with the education community;

3. Geographic proximity;

4. Expertise with open source content management systems.

Experience has shown us that building a strong working relationship with a technology firm is as important as the expertise it brings to the table, so we created a Request for Information (RFI) to help us assess each candidate's technical resources, its ability to grasp the scope and scale of the project, its effectiveness in articulating a solution, and its general approach to working with clients. We sent out the RFI to the ten candidate firms and set a deadline for responding of three weeks, believing that an organization that takes the time to respond in a timely manner is both organizationally sound enough to devote resources to new business acquisition and interested enough to apply those resources to our project.

We received five responses to the RFI. None of the responses met all of the requirements stated in the RFI and all four filtering criteria. The strongest RFI response, for example, came from a firm headquartered in British Columbia, which violated our requirement of geographic proximity. Another strong response was received after the deadline. All, however, presented competitive responses that persuasively articulated organizational capacities, a good understanding of the project and a reasonable plan of attack. We credit the additional step of issuing a RFI—as opposed to simply distributing a Request for Proposal—with ensuring we had a sufficient number of viable candidates from which to choose.

After some discussion, we felt sufficiently impressed by each of the respondents to invite all to submit proposals. While the RFI focused more on organizational capacity and general approach to conducting large-scale technology projects, the Request for Proposal asked candidates to address specific deliverables and provide time and cost estimates. It gave us some measure of confidence that four of the five firms were in the same ballpark when

estimating costs, although one firm disqualified itself by estimating a total project cost at four times that of the others. Though the Canadian firm had a very competitive proposal, we ultimately decided that the geographic distance was just too great; projects of this complexity require a lot of face-to-face contact in order to communicate effectively and build personal relationships. Surprisingly, the strongest candidate coming out of the RFI process submitted a proposal that addressed only half of the RFP's scope, and was the only candidate that misunderstood the requirements of the RFP. This situation was particularly disappointing as the firm in question had extensive experience using our chosen content management system (CMS).

Following a review of the five proposals, which included a checklist to assess a candidate's attention to specific requirements of the RFP and face-to-face meetings, we chose to enter into negotiations with a Boston-based firm that had extensive experience with ECHO-like websites, but that lacked specific experience with our CMS. The firm proposed to bring in outside consultants to address the lack of CMS experience, but we became increasingly uncomfortable with the long-term stability of such an arrangement. We had been burned too many times in the past when a key consultant who was not a permanent staff member of a technology partner left a project to pursue greener pastures. The loss of time and momentum involved in these transitions can be considerable.

We then made a decision that went against the careful vetting process we had followed up to that point: we gave the firm with the great RFI response but weak RFP response another shot. We called the firm's president and had a very frank discussion, describing our initial enthusiasm based on the RFI response and subsequent disappointment with the RFP. We also expressed disappointed in the proposed project team leader, who never seemed to grasp either the nature of the project or our criticisms of the proposal, despite a two hour face-to-face meeting. We offered the firm a chance to submit another proposal provided it included an acceptable replacement for the problematic project team leader.

We found both the resubmission and team member replacement acceptable, and signed a contract with the firm in October for the first phase of the project. The 12-week contract called for the low level design of the site's major landing pages, identification of all significant content types and a navigational scheme. We are just now finishing up that work and have been delighted with both the project team and the work they have done. We very much look forward to working with the group through project launch and beyond.

Some advice we can offer:

1. Make sure the firm you choose has extensive experience with projects very much like yours and the technologies you're using. Make sure they have experience with your cultural environment.

2. Complex projects require frequent face-to-face contact. Skype and email are not enough.

3. If a firm is not responsive or doesn't "get it" upfront, things will only get worse once the contract is signed.

4. If your project has multiple phases, try to get commitment that the core project team will stay in place for the duration.

5. A request for information can weed out unqualified candidates and save time.

6. Indentify your "must haves" and use them as the basis for decision making but don't be too rigid. Judgment is important, too.

7. Personal chemistry is important. It's hard to work on a project with people you don't get along with.

—Brett Thompson

Case Study: Encyclopedia Virginia

Encyclopedia Virginia (EncyclopediaVirginia.org) began as an idea in 2001 when the National Endowment for the Humanities awarded the Virginia Foundation for the Humanities a planning grant to see if creating an online state encyclopedia was feasible. Although the VFH determined that it was well-positioned to take on and accomplish the task of building an online reference work of this scope, it was not until 2005 that the VFH had enough funds to begin the project in earnest. In December 2008, *EV* launched its public website with the priority to have not only authoritative and engaging content but also to leverage the medium in which that content is published to reach as many users as possible through innovative methods.

Goals & Context

Goals, Values, and Mission

When *EV* first began, the intention was to develop a site that would be "Everything Virginia"—a one-stop-shopping of everything about Virginia's history and culture. While this remains a somewhat vague goal of the project, time and resources necessarily limit what a project must prioritize and accomplish. When VFH investigated its capacity to take a project such as *EV* on, they determined the answers to the following:

- Why do we want to do this?

For several reasons and from several perspectives:

- ○ Content perspective: No reference work on Virginia's history has existed since Lyon Tyler's *Encyclopedia of Virginia Biography* (1915) and the WPA's *Guide to the Old Dominion* (1939); need to have a current updateable resource that can communicate a comprehensive and authoritative understanding to the widest possible audience, offering to Virginians and the world a picture of the Commonwealth's history and evolving story.
- ○ Organizational perspective: *EV* could help integrate the disparate "brands" of the entire VFH by promoting content developed and created in other programs. What is more, with the VFH's mission to bring current humanities scholarship into the sphere of public interest, creating an online resource and mechanism to accomplish this mission was important.
- Who needs this?

The key with this question for *EV* as it should be for anyone considering "audience" is determining a very finite and concrete audience to reach. Saying "everyone" becomes quickly counterproductive and dilutes the purpose of a resource. In our case, *EV* focused on two audiences:

- ○ Primary Audience: students and teachers of all ages and all locales interested in Virginia history and culture as part of the national story
- ○ Secondary Audience(s): tourists, general public, scholars
- How will we know it is successful?
  - ○ The general notion of success will be not just with a quantitative number of hits/ unique visits but also with a sense that *EV* is a known go-to resource for users interested in Virginia's history and culture; what is more, success will be defined by how *EV* takes advantage of the growing number of tools and networks that are continually being developed and enhanced in the medium in which it is published.

To help us better define this mission and audience and create a visual brand for it, in 2006 *EV* hired a branding consultant who helped focus our attention on what was most important. This was done as a precursor to developing our web presence and publishing platform (which were also created in 2006-2008). We had to know who we were and who we were trying to reach before instantiating something on the web. This might sound a bit touchy-feely, but

creating an identity while considering both stakeholders and constituents is much easier said than done.

Institutional Partners & Other Relationships

VFH always knew that it would not be able to produce *EV* without a number of partners across the Commonwealth. As Virginia's state humanities council, VFH is not an archive and it does not have its own collections; it does not have the staff to create most of the content and public scholarship necessary to make the site worthwhile; it does not have the infrastructure required to maintain servers on a 24/7 basis. Partnerships, therefore, have been crucial to *EV*'s success. There are several different types of partnerships *EV* has created and maintains either through its own initiative or because of existing relationships with VFH:

- Media Partnerships: in order to feature media (still/moving image and audio) on the site, *EV* creates partnerships with some of the Commonwealth's (and other places) most content-rich organizations. Each relationship is created under terms set forth in a Memorandum of Agreement where *EV* receives high-resolution media as an "in-kind" contribution and the rights statement of those media objects are prominently displayed on the media pages with links back to their source site. Current partnerships include: the Library of Virginia, the Virginia Historical Society, The Museum of the Confederacy, the Special Collections Department at the University of Virginia, Documenting the American South, the Mariner's Museum, and, most recently, The Colonial Williamsburg Foundation.
- Content Partnerships: *EV*, like most projects, works with content contributors (entry writers) from across the state, country, and world. However, *EV* is also working more deliberately with the Library of Virginia to be the publishing platform for its own reference collections including the *Dictionary of Virginia Biography* and the *Hornbook of Virginia History*.
- Infrastructural Partnership: VFH's relationship with the University of Virginia gives *EV* access to the University's server room and server administration expertise to secure content, make backups of the databases and file systems, and provide 24/7 support. This is not a "free" service, but having access to it is a huge benefit to the futurity of the project.

Staff, Users, and Other Stakeholders

EV **Staff**: When the VFH administration made the first hires for *EV*, they decided that instead of getting a content editor to direct the project they needed someone who had a strong background in both the humanities and computing to shape the project in both areas and to be able to communicate between content providers and technologists. Thus what they hired in the managing editor was less a real editor and more of a humanities computing specialist: someone who could understand the needs both of scholars and the general public as users of Internet tools and humanistic content as well as how to translate/implement those requirements in a publishing platform.

Besides a managing editor initial hires for *EV* included an associate and assistant editor whose responsibilities were administering and editing the textual content of the project and a media editor who was responsible for creating partnerships, conducting archival research, and reformatting visual and audio material for the site.

While VFH contracted with an outside firm to help *EV* with branding, web design, and the development of a content management system (CMS), once the core CMS was in place *EV* also hired a full time programmer to carry on with the development of the CMS and further web and mobile applications.

While we do have interns (some paid, some work study) that help with content creation, we prefer to have full time employees as it creates team cohesion and investment in the project. At one time there was an idea that *EV* would hit a "critical mass" of content and that positions might be phased out soon after that mass was reached; however, my perspective is that:

- History changes all the time and to maintain a project's content relevancy (especially on the web), you must stay abreast of new discoveries/research, plus you can always add new content;
- Web technologies are changing all the time and to maintain your *technical* relevance and presence on the web, you must leverage your project as a dynamic tool that can take advantage of new ways to expose your content and build your community of users;
- The task will never be complete and while the project will be difficult to sustain with this philosophy, if a project—especially an encyclopedia—is not thought of in this way, then the product you create will cease to be relevant (after all, there is a reason—besides the profit motive—that Encyclopedia Britannica released new editions every few years).

Works for Hire: Because no one on staff has the content expertise and historical background to develop topics that should be included, *EV* necessarily relied and continues to rely upon works for hire including "section" editors who help create lists of topics (e.g. a list of approximately 350 Civil War entry topics germane to Virginia) with word counts (from 500 words up to 3000 words for more general overview essays) and potential contributors. While some writing of entries occurs in-house, the majority of *EV*'s contributors are also works for hire as are the fact checkers and copy editors who help review content at different stages of the process.

Users: Up to this point, *EV*'s relationship to its users has been relatively passive. We serve them content, they consume it, and maybe they will come back, which we can inaccurately determine through Google Analytics. We have mechanisms whereby the user can comment on content but the comment comes to the Managing Editor as an email. When the user leaves his/her contact information in that email we make an effort to respond in a timely fashion to that individual with appreciation for using the site and for taking the time to comment, etc. We also currently have an email sign up list to which we send an electronic message every month or two that highlights new content and media.

However, lately we have been much more interested in exploiting and leveraging other tools to build a stronger user community that is invested in the work we are producing. Part of this investment has been to appoint our Assistant Editor as the lead on our social media outreach. The person in this position is responsible for being the gateway and proactive communicator through Facebook, Twitter, and the *EV* Blog. While we have automatically produced tweets in the past (and will continue to do this in the future through http://twitter.com/TodayinVa) based on data about what happened today in Virginia, we have realized that a more important use of this network is to be a human being on the other end: to talk about cool stuff going on in with the project, to respond to other topics related to Virginia history and culture specifically or the humanities in general, and to generate conversations with that community of users.

We are also experimenting with a "club-sourced" review of some of our content to build a scholarly investment in the project. By "club sourcing" we post an entry on Googledocs, send the link to a select number of individuals (at least presently a select group, we have ideas of opening this up to a larger crowd if it works) with content expertise to review the entries and offer their knowledge about what is wrong or needs to be enhanced.

Stakeholders: As far as stakeholders, there are many: VFH staff, VFH board, donors (including the Virginia General Assembly and media partners), contributors and editors, the educational community, and the general public user community.

Administration & Management

Governance

EV is subject to the same governing structures that oversee the entire VFH: namely a governing board (as required based on VFH's status as a 501c3) and the University of Virginia

(as VFH is also a department of the University). While the managing editor establishes most of the priorities and procedures of the project, there are organizational processes that are set by VFH and UVa by which *EV* must abide and, of course, there are content creation and technology implementation priorities that funding may also impact.

Finances and Budget

After getting the NEH planning grant in 2001, it took VFH until 2005 to secure enough initial funding from the Virginia General Assembly to begin the project in earnest (noting, of course, that the state and much of the country was going through a heavy shortfall in state funds between 2002 and 2004). However, the Governor Mark Warner was quick to see the potential benefits of an online authoritative resource for both educational and touristic purposes—that an investment in a dynamic work such as *EV* would be worth making. With several hundred thousand dollars allocated to the project for the 2005 fiscal year, *EV* was in a position to hire its first staff and to begin the processes of laying the groundwork to build the technical and workflow infrastructures.

Cost

There will be an up front and possibly hefty cost in both time and money for any project that requires branding, web design, and a content management system (no matter if that system is home grown or out of the box—whether it is open source or proprietary does not really matter, costs will still exist). The initial sticker price for *EV*'s web presence (brand, design, and publication system/content management system) was approximately 200K dollars. The decision to outsource this process was made by VFH's governing board. However, rather than continue to retain the developers of our system for future improvements, tweaks, etc., once we had the foundation of a working code base, we decided that it was in our best interest for the sake of finances and time efficiency to hire our own web programmer. The benefits for us have been great, having a programmer in house:

- Creates a human investment in project and product development
- Provides a better mechanism for collaboration, brainstorming, and innovation in real-time
- Makes *EV* a sustained priority unlike being one of many projects and priorities that external developers have to manage

Since the initial expense of creating the brand, site design, and publishing system, each year that *EV* has been in production, we have budgeted approximately 450K dollars for operations. While the majority of these funds go to pay the full time staff at *EV* (including fringe benefits) as well as the basic fixed costs of infrastructure (rent, computers, server hardware, and maintenance), these monies are also used to pay contributors, section editors, fact checkers, and copy editors for content creation.

"Sustainability"

There seems to be this idea that to have a fiscally sustainable project, a project's power button is flipped "on," some brilliant revenue-generating plan is put in place (while somehow content is still disseminated for free), and then the project runs on a type of autopilot—recouping operational costs as it goes and grows. Perhaps not everyone thinks of sustainability in these terms. That is good. *EV* has always looked at the challenge of sustainability as something that requires a lot of work: a mixture of entrepreneurialism, innovation, and—most important—having a diversification of options.

The methods *EV* employs to maintain its operational budget are diverse. We look to Federal competitive grants as well as corporate and private foundation opportunities. We also rely on individual donors. In 2008, VFH hired an external firm to assist us in a campaign to raise funds for *EV*. Unfortunately, we initiated this endeavor just before the downturn in the economy. However, while the campaign was not as successful as we wished it could have been, it was successful enough to provide for the beginnings of a modest endowment that we have

used to supplement our operational costs.

Currently we are looking at different types of "business models" to buttress at least part of our operational costs. Because the core content—i.e. the publicly accessible web site—will always be free, we are looking to other "added value" formats and ways to aggregate data that might add revenue to the project. Some possibilities include charging a small fee for access to mobile historical points of interest that link back to *EV* content through augmented reality applications such as Layar. This is always a difficult balancing act. We may look to what the Stanford Encyclopedia of Philosophy is doing in regards to its "Friends of …" effort that seems to be somewhat successful and does not compromise access to content.

While salaries are a large part of our budget, one way we are looking to grow the project and its content without encumbering heavy costs is through partnerships and content sharing with other scholarly reference publications (both online and off). Partnerships like this, while not revenue generators, certainly defray costs associated with content production and bring the best and brightest content to bear in one resource. What is more, establishing partnerships with other organizations provides greater leverage in seeking funds from competitive grant programs.

Content

As other case studies have indicated, there are many strategies to produce content. *EV* has primarily chosen to concentrate on particular areas (e.g. a time period or thematic subject) and go deeply into them, produce a critical mass of content, and then move on to another. Regardless of our production method, the big anxiety-ridden question we encountered when we began and one that all projects like this will face was "how *much* content do we need before we can launch; is there a magic number or critical mass we must have before we go live?" I am actually not sure what the answer to this is. In some ways, though, whatever you put out there in whatever quantity should corroborate the philosophy of the brand. If you say you are authoritative, then your entries should reflect that. If you say you provide consistent and accessible content, then your entries should reflect that. I guess the lesson we learned was to be what you say you are as some of our earlier entries reflected a "get it done and get it published" more than a "get it right and consistent" mentality; as far as those earlier entries are concerned, we've had to refurbish some of them or simply pull them off the site until we can find the time to revisit them.

As far as how much content we had when we first launched, we sort of hedged on this. In 2006 we began publishing weekly "Virginia Vignettes" online—snippets of Virginia history that were teasers for what would one day be in *Encyclopedia Virginia*. Soon after this, we started the *EV* blog which helped us connect with potential users by offering them an inside view into what we were working on and why what happened "then" is important "now." An important thing to consider, though, is the amount of time that "sub endeavors" like these can take away from the core project since you never want something like a blog to seem inactive or moribund as that sends a bad message to the community you are trying to build. Overall, we found that the blog has been worth the risk and it has, in fact, helped us build a community.

At the end of 2008, we decided that we should just "go live" in a public beta launch with approximately eighty entries—a far cry from what we wanted to launch with (one thousand entries) when we first started the project. Beta was a great way to get our content out there, build a following of people but also be able to communicate with honesty that there were huge holes in the site's content and that we were still adapting web features for greater usability, etc.

Content Development: Features, Metadata, & Workflow

Features

EV includes the following features with every (or almost every) entry:

- **Entry body:** textual body that covers the subject matter. Each text body includes a "summary" paragraph at the top that attempts to distill the significance of the topic into several sentences. Users can get a sense of what the entry is about by reading the summary and from a technical point of view, we can deploy this paragraph for different needs: e.g. information for our RSS feed, search results on the site, etc.
- **Links:** we maintain links in the body of each entry to other entries (both current and future) and when we "republish all" of our content on a periodic basis, those links get checked and if content exists for them, they are instantiated. We also maintain a "sidebar" section for external links to other pages not controlled by *EV*. There is an overhead to this as it is these links that must be checked on a consistent basis to see if they have moved, etc.
- **Media:** because of our media partnerships, we try and pull in unique visual and audio content where we can and where it is appropriate in the context of each entry. Part of the way we can leverage *EV* as a publishing platform for other archives and collections in the state is that we provide extensive rights information about each media object and always provide links back to the original source: we become a mechanism for giving a higher profile to other organizations' content in a rich narrative context. What is more, almost every media object has its own descriptive information (metadata) that gives the user a deeper understanding of what they are seeing/hearing.
- **Time line:** each entry has a chronological list of events that we cull from the body of the entry and store in a central "master" database table. Because these events are stored as unique items in a separate database table, we can relate and reuse these events in other entries (e.g. chances are that the entry on Pocahontas, John Smith, and Powhatan will all share some similar events) and in other contexts ("Today in Virginia," Twitter, etc.)
- **Mapping:** with the event table mentioned above, we can also relate geographical information to each event if that locale can be determined. Geographic information includes a place name, a place class (monument, river, etc.), and the latitude and longitudinal points for actual mapping. If an entry has this information available, the entry can be "mapped."
- **Further reading & citation:** at the end of every entry we offer a list of further reading and several methods for citing that particular *EV* page.

Metadata

In order to try and leverage the current and future possibilities of content dissemination on the web, *EV* has invested a great deal in its implementation of XML markup standards. While the overall content management system is a relational database/XML hybrid, the core information about our entries and media is created and stored in XML which, because the core content is platform agnostic, allows for easier portability and flexibility for distributing our content through different methods. For our entries, we use an extended version of the Text Encoding Initiative Schema for Markup and Interchange (TEI; www.tei-c.org) and our media metadata uses the Metadata Encoding and Transmission Standard (METS; http://www.loc.gov/standards/mets/) which wraps up several facets of information including descriptive metadata (using qualified Dublin Core) and technical and rights metadata. While this level of markup requires effort and time, the payoff is in the fact that we need to spend only small amounts of time writing programs to give users different views of our content through different contexts and mediums whether it is a geographic and temporal visualization, creating dynamic print-on-demand aggregations of content, eBooks, etc., these are easier to do when the core content is built well from the beginning. *EV* abides by the mantra: "build content once, use it many ways."

Workflow, Intellectual Integrity, & Change Management

EV follows much the same workflow and quality control model as other reference works in this environment. We hire a section editor(s) (someone who has competence with a particular section of material and is well-networked with scholars/experts in the area—see "Works for Hire" above). The section editor's tasks are:

1. To generate a list of topics in the particular thematic area they are charged to curate.

2. To add word counts (between 500 and 3000 words for individual pieces) and contact info for possible contributors for those pieces.

3. To vet content when it comes in (after EV staff perform a range of triage on those pieces) to ensure that the topic is covered and addresses what must be addressed.

The list itself is peer-reviewed by other scholars in the field and when it is finalized, *EV* staff begins the job of recruiting writers. If the writer accepts our offer, we provide a contract and schedule A (this identifies deadlines, word length, and honorarium), "Contributor Guidelines" [see attachment A], and one or more example entries.

Once we receive the entry from the contributor, *EV* editors perform varying amounts of triage on it before sending it to the section editor for further review (e.g. checking comprehensiveness of entry and general accuracy of information).

Once we receive the entry from the section editor with comments, we resolve any issues that the section editor found by:
 • asking the original contributor for clarification;
 • or, and what we find to be more efficient, doing the research ourselves

Once we have a clean copy, the entry goes to fact checkers and then the iterative process between comment and resolution begins again; the same thing happens with the final process of copyediting.

Once we have a final copy, we prepare it for XMLification [see attachment B]. We extract time and place information from the original entry, store that information in our database, create internal links to other entries, and add media to the entry if it is available at this time. We publish the entry to a staging site in order to copyedit the online entry, make emendations to the XML as we see fit, and then publish the entry to our live site.

Even though the entry is finished at this point, *EV* believes that its content (like any reference content on the web) has got to be maintained and updated and that while some content might be more "finished" than others, nothing, to paraphrase Quentin Compson's ruminations, is ever really finished. It may seem that this state of being is truer with works on the web than those that are in print, but it isn't; it is just that the feedback loop and a user's expectations on the web and with web content are different. Right now the way users report errors of fact, interpretation, or syntax is through a "feedback" button that appears on each entry. This feedback mechanism is somewhat effective and EV staff is always responsive to users if they provide contact information. At times we turn to further primary document research to clear up facts or at least the facts as we know them. There are also times when we use our Blog to talk about a recent factual error and how we deal with correcting it or proving that what we have is correct—this is a nice way to provide some transparency in how EV does things and why. When content has been altered, we publish a "most recently published" date at the bottom of the entry and in its citation. We also have an "originally published" date to show our users that the entry has evolved over time; however, while this is not a priority it may be potentially useful to link users back to previous versions of an entry.

—Matthew Gibson

Attachment A

Encyclopedia Virginia Contributor Guidelines

General guidelines

1. Make your writing clear and accessible. Define subject-specific terms where appropriate. No need to dumb down the prose—just assume that the reader knows nothing about the subject.

2. Keep the pace brisk and the tone encyclopedic, but write with an eye for the telling detail or anecdote.

3. Do not shy away from controversy, debate, unanswered questions, etc. You are invited to make arguments; however, frame such discussions so that any arguments are clearly marked as such. Make sure the reader fully understands the historical context and is able to distinguish between what is the consensus view of historians and what is not.

4. Write biographical entries in chronological order (i.e., after the introduction, begin with the subject's birth and end with his or her death). When writing about a person, do not skip major events in your subject's life; if said events have no relation to Virginia, they can be quickly summarized.

5. If you need to write longer than the assigned word count, please do. Maximum: 3,000 words.

6. Attribute all quotations to their speaker and/or source. If you are unable to do so in the text, write a note in the margin. In the margin, provide a source for any material (e.g., unpublished primary sources) that would not be readily available for our fact checkers.

7. When quoting, always use original (mis)spellings.

8. When mentioning publications in the text, provide the full title and publication year.

9. Use as many exact dates as possible (i.e., day, month, year). If such dates clog your prose, then include them in the time line.

10. Finally, please look over entries on the *EV* site. These are your models.

When writing about a campaign or battle

1. Provide the following pieces of data, which will be placed in a sidebar: a) name of campaign or battle; b) date and place; c) opposing commanders; d) casualties for both sides, including killed, wounded, and missing (if known). This information should be considered separate and may be repeated in the narrative portion of the entry.

2. Organize your entry into three major subheads: 1) "Background"; 2) "The Battle" (or "The Campaign," as appropriate); and 3) "Aftermath."

Time lines

Create a time line out of all the dates mentioned in the entry (and any important dates that did not otherwise make the cut). Include 1) Specific date; 2) Specific place; and 3) Description of event. Write in the present tense. The time line is **not** figured into your assignment's word count.

time line Excerpt for robert bolling

- **August 17, 1738:** Robert Bolling is born in Henrico County. He is a great-great-grandson of Pocahontas and John Rolfe.
- **1751:** Robert Bolling's father, John Bolling, sends him to the Free Grammar School of Queen Elizabeth in Wakefield, England, which a number of other Virginia boys attend. Bolling excels in languages.
- **December 31, 1755:** Robert Bolling is admitted to the Middle Temple in London, England, for legal studies.
- **April 16, 1756:** After studying law in London, England, Robert Bolling returns to Virginia, arriving in Yorktown. He continues study of the law for a year and a half under Benjamin Waller in Williamsburg.
- **1757:** Robert Bolling's father, John Bolling, dies. Bolling lives at his father's Chesterfield County plantation, Cobbs, until the middle of 1760, when he builds his own home in southern Albemarle County.
- **January 12–September 16, 1760:** Robert Bolling courts Anne Miller, a distant cousin. He keeps a courtship journal, "A Circumstantial Account," which provides a unique view of eighteenth-century Virginia gentry. The romance ends when Miller sails for Scotland with her father.
- **1761–1765:** Robert Bolling is elected to and attends the assembly of the House of Burgesses.

Related Websites

Please recommend existing Internet sites (beyond *EV*) that provide further information for readers on the topic at hand. Give the basic URL for a site (http://www.vahistorical.org) or, when appropriate, a specific page within a site (http://www.vahistorical.org/civilrights/main.htm). Provide the name of the site, too (Virginia Historical Society, http://www.vahistorical.org). Through these references, users can explore other important resources outside *EV*. Be selective and discerning in what you recommend.

Further Reading, Listening, and Viewing

Provide a "Further Reading" list comprising the most important and up-to-date print, video, or audio sources that readers might consult for further study of the topic.

Originality

Entries for *EV* must be original work. Please avoid repeating or following too closely the content of other articles on the same topic, particularly other Internet materials. Contributors are responsible for any violation of copyright. Plagiarism of any kind will result in the contract with the writer being nullified.

Bibliography

Please include a list of the sources used in the creation of the entry. While this bibliography will not be published with the entry, it will be referred to by fact checkers and editors if questions arise during the editing process.

biographical note

Please provide a brief one- to two-sentence description of yourself that begins with your name and includes any institutional affiliations and appropriate cities/states. If you include a published a work, place the publication year in parentheses.

Sample Entry XML

```xml
<?xml version="1.0" encoding="UTF-8"?><?oxygen SCHSchema="../tools/evtei.
rng"?><?oxygen RNGSchema="../tools/evtei.rng" type="xml"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0" xml:id="LostCause">
  <teiHeader>
   <fileDesc>
     <titleStmt>
       <title>Lost Cause, The</title>
       <respStmt>
         <resp>Contributor</resp>
         <name key="307"/>
       </respStmt>
       <respStmt>
         <resp>Section Editor</resp>
         <name key="196"/>
       </respStmt>
       <respStmt>
         <resp>Editor</resp>
         <name key="326"/>
       </respStmt>
       <respStmt>
         <resp>Conversion to TEI-conformant markup with EV local extension</resp>
         <persName>
           <forename>Matthew</forename>
           <surname>Gibson</surname>
         </persName>
         <orgName>Virginia Foundation for the Humanities</orgName>
       </respStmt>
     </titleStmt>
     <publicationStmt>
       <publisher>Encyclopedia Virginia at the Virginia Foundation for the
         Humanities</publisher>
       <pubPlace>Charlottesville, VA</pubPlace>
       <idno type="EV">LostCause</idno>
       <availability status="publiclyaccessible" url="http://www.encyclopediavirginia.org">
     <p>This entry is publicly accessible; however, individuals interested in reproducing the
text of this entry in a publication, web site, or for any commercial purpose must first receive
written permission from Encyclopdia Virginia (info at encyclopediavirginia dot org).</p>
       <p>Permissions regarding images and multimedia files are restricted and individuals
wishing to use those materials must contact the original owner(s).</p>
       </availability>
       <date>2008</date>
     </publicationStmt>
     <seriesStmt>
       <p>This entry is part of Encyclopedia Virginia.</p>
     </seriesStmt>
     </fileDesc>
   <encodingDesc>
     <projectDesc>
   <p>Prepared for Encyclopedia Virginia at the Virginia Foundation for the Humanities</p>
     </projectDesc>
     <editorialDecl>
      <p>Spell-check and verification made against Merriam Webster's Collegiate Dictionary
11th Edition<p>
<p>When possible, images exist as archived tiff images with one or more jpeg versions for
```

> These 'keys' link off to a table in our database that contains information about each contributor (e.g. biography, etc.)

```
general use.<p>
<p xml:id="EV">Keywords in the header are a local Encyclopedia Virginia scheme to aid in
establishing analytical groupings and cross references.</p>
        <p>All place and feature localities are keyed to USGS Feature ID codes.</p>
    </editorialDecl>
    <classDecl>
      <taxonomy type="ev">
        <category>Civil War, American (1861–1865)</category>
        <category>Memorials and Commemorations</category>
        <category>Reconstruction and the New South (1865–1901)</category>
        <category>Twentieth Century History (1901–2000)</category>
        <category>African American History</category>
        <category>Political Issues and Controversies</category>
        <category>Women's History</category>
      </taxonomy>
    </classDecl>
  </encodingDesc>
  <profileDesc>
    <creation>
      <date>2009–07–07T10:28:00</date>
    </creation>
    <settingDesc>
      <place>
        <geogName key="2092203"/>
        <geogName key="2090198"/>
        <geogName key="2092170"/>
        <geogName key="2092213"/>
        <geogName key="2092214"/>
        <geogName key="2092215"/>
        <geogName key="2092216"/>
        <geogName key="2092217"/>
        <geogName key="2092218"/>
        <geogName key="2092219"/>
        <geogName key="2092220"/>
        <geogName key="2092221"/>
        <geogName key="2092222"/>
        <geogName key="2092223"/>
        <geogName key="2092224"/>
        <geogName key="2092225"/>
        <geogName key="2092226"/>
        <geogName key="2092227"/>
        <geogName key="2092228"/>
        <geogName key="2092229"/>
        <geogName key="2092230"/>
        <geogName key="2092231"/>
        <geogName key="2092232"/>
      </place>
    </settingDesc>
    <chronList>
      <event key="cl3069"/>
      <event key="cl3070"/>
      <event key="cl3071"/>
      <event key="cl3072"/>
      <event key="cl3073"/>
      <event key="cl3074"/>
```

> These 'keys' link off to a table in our database that contains information about each geographical point lists (name, feature type, lat/long)

```
<event key="cl3075"/>
<event key="cl3076"/>
<event key="cl3077"/>
<event key="cl3078"/>
<event key="cl3079"/>
<event key="cl3080"/>
<event key="cl3649"/>
</chronList>
<langUsage>
  <language ident="eng">English</language>
</langUsage>
<textClass>
  <keywords scheme="cdata">
<term>Lost Cause, slavery, secession, states' rights, faithful slaves, Robert E. Lee, Mary
Tucker Magill, Thomas Nelson Page, Gone with the Wind, Pickett's Charge, Cold Harbor,
James Longstreet, Jubal A. Early, Douglas Southall Freeman, Thomas J. "Stonewall" Jackson,
John Esten Cooke, Robert Lewis Dabney, Gettysburg, Richmond
Bread Riot, Edward A. Pollard, Ladies' Memorial Associations, H. Rives Pollard,
Southern Historical Society, United Confederate Veterans, Sons of Confederate
Veterans, United Daughters of the Confederacy, LaSalle Corbell Pickett,
Confederate battle flag, Frederick Douglass</term>
  </keywords>
</textClass>
</profileDesc>
<revisionDesc>
<change who="MSG" when="2009–07–07T10:28:00">Converted original document
(MS Word or RTF) to TEI P5 document conformant document using Open Office and
XSLTAdded TEI header and base entry tags</change>
<change who="PL" when="2009–07–07T10:28:00">Checked entry markup and conformed
to TEI-EV schema specifications. Added divison headings and structure.</change>
  <change who="PH" when="2009–07–07T10:28:00">Embedded media references for all
media content</change>
  <change who="MSG" when="2010-10-08T13:54:12.392-04:00">Updated markup to
conform to new schema.</change>
</revisionDesc>
</teiHeader>
<text>
  <body>
    <head>The Lost Cause</head>
    <div type="entry">
      <p type="summary">
<figure url="evm00002464" rend="right"/>The Lost Cause is an interpretation of the
American Civil War (1861–1865) that seeks to present the war, from the perspective of
Confederates, in the best possible terms. Developed by white Southerners, many of them
former Confederate generals, in a postwar climate of economic, racial, and
 gender uncertainty, the Lost Cause created and romanticized the "Old South" and the
Confederate war effort, often distorting history in the process. For this reason, many
historians have labeled the Lost Cause a myth or a legend. It is certainly an important
example of public memory, one in which nostalgia for the Confederate past is accompanied
by a collective forgetting of the horrors of <ref xml:base="entries" url="Slavery.
xml">slavery</ref>. Providing a sense of relief to white Southerners who feared being
dishonored by defeat, the Lost Cause was largely accepted in the years following the war by
white Americans who found it to be a useful tool in reconciling North and South. The Lost
Cause has lost much of its academic support but continues to be an important part of how the
Civil War is commemorated in the South and remembered in American popular culture.</p>
```

These 'keys' link off to a table in our database that contains information about temporal events related to the entry (e.g. description of event, year/month/day, etc.)

<div type="section">
  <head>Six Tenets</head>
    <p>The Lost Cause interpretation of the Civil War typically includes the following six assertions: <list>
<item><ref xml:base="entries" url="ConstitutionalConvention1861.xml">Secession</ref>, not slavery, <ref xml:base="entries" url="CausesCivilWar.xml">caused the Civil War</ref>.</item>
      <item>African Americans were "faithful slaves," loyal to their masters and the Confederate cause and unprepared for the responsibilities of freedom.</item>
        <item>The Confederacy was defeated militarily only because of the Union's overwhelming advantages in men and resources.</item>
        <item>Confederate soldiers were heroic and saintly.</item>
<item>The most heroic and saintly of all Confederates, perhaps of all Americans, was <ref xml:base="entries" url="LeeRobertE.xml">Robert E. Lee</ref>.</item>
<item><ref xml:base="entries" url="WomenCW.xml">Southern women</ref> were loyal to the Confederate cause and sanctified by the sacrifice of their love ones.</item></list> The historical consensus, however, presents a picture that is far more complicated, one in which some tenets of the Lost Cause are obviously false and some are at least partly true.</p>
    <p><figure url="evm00002293" rend="left"/>Lost Cause proponents have stressed the primacy of <ref xml:base="entries" url="StatesRights.xml">states' rights</ref> and the constitutionality of secession, and have cited the secession crisis—along with political squabbles such as tariff disputes and broad claims about the evolution of different societies in the North and South—as the cause of the war instead of slavery. At the same time, Northern abolitionists have been portrayed as provocateurs and slavery as justified in part as an institution that eventually would have died of its own accord. The historian Alan T. Nolan has called this reading of history "outrageous and disingenuous," suggesting that it was the dispute over slavery that actually caused the secession crisis. Nolan and other historians have further noted that many Southern politicians viewed slavery to be, in the words of Confederate vice president Alexander H. Stephens, the "foundation" and "cornerstone" of the Confederacy.</p>
    <p><figure url="evm00002262" rend="right"/>Slavery, meanwhile, is sentimentalized in the context of the Lost Cause. Following the war, white Southerners told stories of the happy slave, the "Mammy" or "Uncle Tom" who appeared as part of the family. "Generally speaking, the negroes proved a harmless and affectionate race, easily governed, and happy in their condition," according to the 1908 edition of the textbook <hi rend="italic">History of Virginia</hi> by <ref xml:base="entries" url="MagillMary.xml">Mary Tucker Magill</ref>. The 1964 edition of <hi rend="italic">Virginia: History, Government, Geography</hi> by Francis Butler Simkins, Spotswood Hunnicutt Jones, and Sidman P. Poole was not much different. "A feeling of strong affection existed between masters and slaves in a majority of Virginia homes," the authors wrote. Such statements are not supported by modern scholarship, which suggests that many slaves were desperate to escape their often harsh conditions both before and during the war, when they became <ref xml:base="entries" url="RefugeesCW.xml">refugees</ref>. In fact, escaped slaves helped to precipitate national political crises such as the one surrounding the <ref xml:base="entries" url="FugitiveSlaveAct.xml">Fugitive Slave Act</ref> of 1850.</p>
    <p>The image of African Americans who had been happy under slavery but were overwhelmed by the responsibilities of freedom became widespread and could be found in the fiction of <ref xml:base="entries" url="PageTN.xml">Thomas Nelson Page</ref> and Margaret Mitchell, whose novel <hi rend="italic">Gone with the Wind</hi> won the Pulitzer Prize in 1937. The image also proved particularly useful to white supremacists. In the 1880s and 1890s, white Southerners, decrying "Yankee aggression" and black "betrayal," embarked on an effort to reverse the policies of Reconstruction (1865–1877). They sought to remove black office holders, <ref xml:base="entries" url="Disfranchisement.xml">disenfranchise</ref> African American men, forestall black economic advancement, and institute state-sanctioned segregation.</p>

[…]

```
        </div>
    </div>
        <div type="externalresources">
          <p>
            <ref url="http://www.moc.org/">Museum of the Confederacy</ref>
            <ref url="http://www.lib.lsu.edu/special/findaid/u1357.html">United Confederate
Veterans Records</ref>
            <ref url="http://www.scv.org/">Sons of Confederate Veterans</ref>
            <ref url="http://www.hqudc.org/">United Daughters of the Confederacy</ref>
          </p>
        </div>
        <div type="bibl">
          <list>
            <item>Blight, David W. <hi rend="italic">Race and Reunion: The Civil War in
American Memory</hi>. Cambridge, Massachusetts: Harvard University Press, 2001.</item>
            <item>Connelly, Thomas L. <hi rend="italic">The Marble Man: Robert E. Lee and
His Image in American Society</hi>. Baton Rouge: Louisiana State University Press, 1977.</
item>
            <item>Cox, Karen L. <hi rend="italic">Dixie's Daughters: The United Daughters of
the Confederacy and the Preservation of Confederate Culture</hi>. Gainesville: University
Press of Florida, 2003.</item>
            <item>Foster, Gaines M. <hi rend="italic">Ghosts of the Confederacy: Defeat, the
Lost Cause, and the Emergence of the New South.</hi> New York: Oxford University Press,
1987.</item>
            <item>Gallagher, Gary W. <hi rend="italic">Lee and His Generals in War and
Memory</hi>. Baton Rouge: Louisiana State University Press, 1998.</item>
            <item>Gallagher, Gary W. and Alan T. Nolan, eds. <hi rend="italic">The Myth of
the Lost Cause and Civil War History</hi>. Bloomington: Indiana University Press, 2000.</
item>
            <item>Janney, Caroline E. <hi rend="italic">Burying the Dead but Not the Past:
Ladies' Memorial Associations and the Lost Cause</hi>. Chapel Hill: University of North
Carolina Press, 2008.</item>
          </list>
        </div>
      </body>
    </text>
</TEI>
```

e-WV: The West Virginia Encyclopedia

The West Virginia Humanities Council launched its digital encyclopedia, *e-WV: The West Virginia*, in September 2010. The project traces its roots back to 1997 when the Humanities Council began work on the West Virginia Reference Project. The council worked for nearly a decade on a print encyclopedia called the *West Virginia Encyclopedia*, which includes more than 2,200 articles by nearly 600 writers. This vast resource published in 2006 became the initial knowledge base for *e-WV*.

As the West Virginia Humanities Council began planning for a digital encyclopedia, it identified the critical technical considerations that pertained to system design. The Council met with two types of advisers during this period: technical individuals who could comment on the core technology available and experts in other fields (education, history, etc.) who could suggest valuable features outside the basic encyclopedia references. The Council also met with groups who had already published an online encyclopedia to learn from their experiences.

The Council hired a software developer, Information Research Corporation of Fairmont, West Virginia, that designed two custom web applications--a public website and a content management system (CMS). It was imperative that the system be designed on an open architecture which could be maintained and upgraded by IT staff outside the initial contractor. IRC met with the council at least monthly for 18 months to complete the work, and as the website became functional, Council staff provided feedback and suggestions to improve both the public and private parts of the site from a user point of view. The West Virginia Humanities Council maintains a relationship with this company to provide technical support when needed. The website is hosted offsite by another West Virginia company.

 The West Virginia Humanities Council hired two additional staff members who work full-time on *e-WV*. They make updates to articles to keep them relevant and continue to improve the content of the site with the addition of photos, audio and video. They respond to users who suggest corrections to articles and additional material for *e-WV*, and they promote the site, especially among educators and librarians. In addition to the full-time staff members, the council also hires freelance writers and other experts who help improve the content of *e-WV*.

The Council provided the necessary start-up funds for *e-WV*. The project also was funded through a major grant from Verizon. The council continues to seek additional funding for the project.

—Becky Calwell

Handbook of Texas Online

1. Goals & Context

The Handbook of Texas Online was developed as a print to digital conversion based on the 1996 *New Handbook of Texas*, a six volume revision of the Handbook of Texas, originally published in two volumes in 1952, with a third volume added in 1976. The revision project began in 1982 with publication of the NHOT in 1996. Digital conversion began in 1997 and the online Handbook was launched in 1999. For the purposes of this paper, the two products—NHOT and Online HOT—can be seen as two outputs of one project and they are treated as an integrated project in this case study.

A. Purpose, Values, & Mission

The stated purpose for the project remained grounded in the objectives that guided the original *Handbook*—producing a concise, authoritative, accessible work that covered the full range of Texas history. In addition to updating and correcting existing entries, and adding new entries on subjects that had arisen since publication of the previous volumes, the Association aimed to expand the scope of entries to include subject areas that had received new, or renewed, attention including: business history, cultural and ethnic history, environmental subjects, and the history of women, among others. The values guiding the *Handbook* reflected TSHA's traditional focus on the dual values of academic scholarship and lay participation. So, while selection, production, and review of entries was framed by traditional academic structures, the scope and specificity of entries (e.g., the inclusion of more than 5,000 community entries) reflected the Association's interest in having the *Handbook* serve the broadest possible audience.

B. Institutional Partners & Other Relationships

As had been true from the beginning of the first *Handbook* effort, the University of Texas served as the Association's primary partner in the new effort, providing housing, logistical and technical support, and furnishing some of the necessary editorial salaries. Participation was by no means limited to the University of Texas, however. Institutions of higher education throughout the state, as well as regional and local historical societies, served as co-sponsors, provided staff support and funding, and encouraged their graduate students to participate in the project.

C. Staff, Users, & Other Stakeholders

Staff. The core staff consisted of three components: a full-time editorial staff reporting through a managing editor to the Editor-in-Chief, a combination of part-time/full-time staff of graduate student/post-graduate writers, and a part-time graduate student staff of fact checkers. Volunteer staff included 60 advisory editors, several hundred reviewers, and approximately 1,000 contributing authors.

Users. Potential users for the *Handbook of Texas* range from 4th grade students (in Texas, state history is covered in the curriculum at 4th, 7th, and high school levels) to post graduate researchers and include K-12 teachers, college undergraduate and graduate students, academic and independent scholarly researchers, genealogists, local history enthusiasts, government officials, journalists, travel agents, among others.

Stakeholders. Several groups have significant vested interests in the *Handbook of Texas*, starting with the membership and Board of Directors of the TSHA (the *Handbook*'s owner) and extending to senior managers at partner institutions such as the University of Texas and the University of North Texas, K-12 and university level educators throughout the state.

2. Administration & Management

Administration and management of the *Handbook* has always been provided by the staff and Board of Directors of the TSHA.

A. Governance

Final authority for budgetary aspects of the *Handbook* program rests with the TSHA Board of Directors, who must approve annual budgets for all TSHA programs. Executive authority for editorial and operational issues is vested in the association's Executive Director, who serves as the *Handbook*'s Editor-in-Chief.

B. Funding & Budget

Development of the *New Handbook of Texas* was funded through a combination of financial contributions from members and charitable foundations, NEH grants, and in-kind support provided by institutions of higher education, especially the University of Texas. Annual budgets ranged from just under $100,000 to approximately $500,000 at the height of the *Handbook*'s staffing in the early 1990s.

C. Communication

Primary communications with stakeholders during the revision process were handled through the TSHA's official publications—the *Southwestern Historical Quarterly* and *Riding Line*, a quarterly newsletter. Numerous articles about the project appeared in newspapers around the state, particularly as the publication date approached. Members of the project staff regularly made presentations at meetings of historical organizations in the state, as well as key national meetings, primarily those of the American Association for State and Local History, the Southern History Association, and the Western History Association. Each of these methods has been continued since launch of the online Handbook, but the primary modes of communication have shifted to electronic mechanisms via the project web site.

D. Rights & Licensing

Contributing authors for the *New Handbook of Texas* signed a contract conveying copyright to the Texas State Historical Association, which holds the copyright to the publication. The limited number of illustrations included in the six-volume edition were licensed from their owners through the services of an illustrations editor with assistance from contract illustrations researchers. Contributors to the *Handbook* subsequent to publication of the online *Handbook* agree to provide the Association with a non-exclusive license to use their work in all versions of the *Handbook*, both online and print. Media rights are negotiated individually with the copyright holders.

3. Content

Content development for the *Handbook of Texas* began in 1940 with launch of the initial print project. The 18,000 articles developed for that 1952 publication formed the base from which the subsequent print and online editions were developed. Many of those original articles, of course, were either substantially revised or replace outright in subsequent editions, especially the 1996 *New Handbook of Texas.*

A. Content Development

Articles for the *New Handbook of Texas* were developed from a table of contents developed jointly by the editorial staff and volunteer advisory editors for some sixty topical and chronological subject areas. With few exceptions, these advisory editors were drawn from senior ranks of academic historians at institutions throughout Texas and the United States,

reflecting the Association's traditional emphasis on scholarly rigor. Topic lists developed by staff and advisory editors, were augmented with suggestions solicited from numerous user groups including county historical commissions, K-12 educators, librarians at all levels, and members of regional historical organizations.

Similarly, authors for most substantive historical articles were recruited from the ranks of published authors, based on recommendations of advisory editors and other scholarly advisors. The volume (24,000) and extreme range (thousands of very small communities and geographic features, and thousands of biographical sketches), however, required that authors be recruited from a variety of directions, including graduate students enrolled in special courses at participating universities, genealogical and community history enthusiasts, and broad range of individuals having specialized knowledge of specific topics. Finally, late-stage graduate students, including some post-doctoral fellows, were hired as dedicated writers to augment areas where the volume of entries exceeded available volunteer authors.

B. Quality & Accessibility

Submitted articles went through several levels of review. Advisory editors reviewed the most significant articles in their subject areas for completeness and general factual accuracy. These reviews were augmented in many cases with peer reviews by specialists in the subject area. All articles were fact-checked by the *Handbook*'s staff of research assistants using footnotes provided by the authors and drawing upon the massive resources of the now Briscoe Center for American History at the University of Texas.

Handbook editors aim for a uniform style and reading level across entries while, at the same time, preserving the individual voices of the thousands of authors as much as possible. To facilitate this objective, all authors were provided with customized guidelines for articles in their subject areas that provided them with a general outline of desired content and advised them on stylistic guidelines. To insure a reasonable level of stylistic and expressive consistency across entries, full-time text editors served on staff throughout the project, peaking at three text editors during the peak years of content development.

C. Metadata

Metadata as such was developed in only a limited fashion for the *New Handbook of Texas*. Topical, geographical, and chronological codes were developed and recorded in data files about each entry in order to generate specific lists of topics for advisors and potential authors interested in entries related to particular subjects, locations, or time periods. In addition, editors carried over an internal notation system (superscripted qv and qqv notations which indicated the presence elsewhere in the encyclopedia of an article carrying the title to which the qv notation was applied) that guides readers to other articles in the encyclopedia, as well as the more standard see also references at the end of articles. With development of the online edition, these qv notations have been converted to hyperlinks providing even more ease of access for the online user. The content of the Handbook has recently been moved into a Drupal-based content management system, and the Association is currently evaluating strategies for developing full-fledged xml metadata to expand the versatility of the encyclopedia content.

D. Archiving

During the course of the print project, electronic files were manually backed up locally on a daily basis, and were backed up to an offsite location on a weekly basis. The online edition is automatically backed up to offsite servers on a daily basis. In addition, the hard copy working files pertaining to all entries are available back to 1940.

4. Infrastructure

The TSHA has always partnered with a leading institution of higher education for infrastructure support. For the first 100+ years of its existence, infrastructure support was provided by the University of Texas. In 2008, the Association moved to facilities on the campus of the University of North Texas and began a new institutional relationship. In each case, the partner institution has been instrumental in providing physical infrastructure (housing, utilities) as well as Information Technology support, and the Association has generally patterned its hardware and software development plans in consultation with its University partner.

—Doug Barnett

Case Study: New Georgia Encyclopedia

The New Georgia Encyclopedia (**www.georgiaencyclopedia.org**) launched in February 2004 with nearly 700 articles. Today, the NGE offers 2,100 articles and more than 6,200 images and video and audio clips. The NGE has been recognized with several awards, including the Helen and Martin Schwartz Prize for Public Humanities Programs (twice); the Leadership in History Award, given by the American Association for State and Local History; the History in the Media Award, given by the Georgia Historical Society; and the Best Reference Source on the Web, awarded by *Library Journal*. The website averages 1 million page views per month, and is heavily used in higher education and K-12 institutions in Georgia.

1. Governance

Creating a Partnership

The New Georgia Encyclopedia was developed through a partnership of the Georgia Humanities Council, the Office of the Governor of Georgia, the University of Georgia Press, and the University System of Georgia/GALILEO.

The initiative grew out of a partnership forged in the mid-1990s when the Georgia Humanities Council and the University of Georgia Press jointly published *The New Georgia Guide*. Five years in the making, the *Guide* was a monumental effort that involved hundreds of individuals in the state in the planning, fundraising, researching, and writing. The result was an accurate and thoughtful portrayal of the state and an outstanding example of what can be accomplished through bringing together people and resources from state government, the university system, and the Georgia Humanities Council, with support from the private and public sectors.

Seeking to build on the success of that collaboration, in 1998 the Georgia Humanities Council and the University of Georgia Press convened a group of the state's leading scholars, archivists, educators, and policymakers to explore the feasibility of developing an encyclopedia for Georgia—a comprehensive reference work that would document the state's history, culture, and resources. Those attending the meeting strongly and enthusiastically supported the initiative. The discussion moved rapidly to an exciting vision for a reference work that would be of interest and utility to the citizens of the state, policy and decision makers, institutions of learning at all levels, libraries, and others interested in Georgia. The project was seen as an opportunity to produce much more than a reference book and to break ground with new technology. The name, the New Georgia Encyclopedia, pays homage to *The New Georgia Guide*, but it is "new" in a much larger sense. As an online publication the encyclopedia could be continuously updated and expanded and thus be perpetually new.

Following the meeting, the Council and the Press presented a preliminary proposal for the project to then Governor Zell Miller, who had commissioned *The New Georgia Guide*. Governor Miller committed $100,000 in state funds for the initial phase of planning and development. A planning committee met regularly over the next several months to outline a blueprint for the project and fundraising strategies.

The encyclopedia was initially proposed as both a print volume and a Web-based multimedia electronic publication. As the planning process unfolded, however, the clear consensus among the project partners was that a reference work of this scope and magnitude could best be published in an electronic format and should be developed as such from the outset. The proposed encyclopedia would manifest the same editorial standards and integrity as traditional print publications from the University of Georgia Press, the scholarly publisher whose imprint the electronic encyclopedia would bear.

In fall 1998 GALILEO, an initiative of the Board of Regents of the University System of

Georgia, joined the Press and the Council as a project partner. As one of the nation's first virtual libraries, GALILEO (Georgia Library Learning Online) was a natural third partner. Launched in 1995, GALILEO began as a database for the institutions in the University System of Georgia to share information resources. Over the next five years, as funding became available, private colleges and universities within the state, public libraries and schools, and private citizens also gained access. Additionally GALILEO (in partnership with the University of Georgia Libraries) was creating the Digital Library of Georgia, a project in which historical documents and otherwise unavailable materials are digitized so that they may be accessed online. As an NGE partner, GALILEO offered the perfect "home" for the online encyclopedia as well as professional expertise, consulting, and high-level technical support.

In early 1999 the planning committee engaged Merrill-Hall New Media to design and build the information architecture to support the online encyclopedia. Merrill-Hall produced a database, using eXcelon, and a text editor ("publishing tool") for adding content.

In fall 1999 and spring 2000, the Georgia Humanities Council spearheaded an intensive fundraising effort for the project. Governor Roy Barnes, who followed Zell Miller as the state's chief executive, committed an additional $400,000 in state support to the project early in his new term. Funding commitments were also secured from the Robert W. Woodruff Foundation, Georgia Power Foundation, BellSouth, Peyton Anderson Foundation, James Cox Foundation, UPS Foundation, Historic Chattahoochee Commission, and the original project partners, who made substantial in-kind commitments.

The Office of the Governor joined the initiative as a fourth partner, with Governor Barnes agreeing to serve as honorary chair of the project and former Governor Zell Miller agreeing to serve as honorary co-chair. Governor Barnes subsequently appointed an Advisory Board, whose role was to offer advice, guidance, and recommendations to the project's Executive Board. The board included institutional as well as individual (at-large) members, who served three-year terms.

Meanwhile, in July 1999 Dr. John Inscoe, professor of history at the University of Georgia and then editor of the *Georgia Historical Quarterly*, was hired as the project's editor. In November Dr. Nancy Grayson, formerly of the University of Georgia Press and the University Press of Kentucky, was hired as project coordinator, and the project officially opened its editorial offices, housed at the University of Georgia in Athens. Soon, an assistant editor, media editor, editorial assistant, managing editor, and project editor were hired.

From 1999 to 2003 the budget for the New Georgia Encyclopedia totaled $2.4 million dollars, with $1.9 million in cash and $588,000 in in-kind contributions from the project partners. By 2010 the annual budget for the project was approximately $290,000.

Governing the Project

Partners agreed at the outset to divide responsibilities for different aspects of the project; no single organization or figurehead could claim total responsibility. This agreement was a mutual understanding, without any formal documentation, and the responsibilities, or in-kind contributions, were deemed to be more or less evenly divided. Thus, the Georgia Humanities Council handled fundraising; GALILEO handled technical support; NGE staff, working independently in office space donated by the University of Georgia but under the auspices of the University of Georgia Press, handled content development; the Press handled most administrative tasks, including NGE staff salary disbursement, and some marketing and promotion efforts; and the Office of the Governor ensured state support and funding for the project. An executive committee--still active today--includes members from each project partner and serves an oversight function.

The success of the NGE was far from assured. With little idea of what to expect once the website launched, the partners had no concrete criteria for determining success other than "reaching" the target audiences. It was difficult to know how long project staff would be required and in what capacity, as well as whether and how long funding would be available.

There were no formal agreements among the partners regarding the length and extent of commitments. There were also no formal protocols or procedures for decision-making by the partners (via their participation on the executive committee) and the editorial staff; no chain of command was established either within the executive committee or between the committee and the project staff. Occasionally this structure has resulted in a slower decision-making process.

Shifting Demands upon Partners

The Council soon emerged as the "lead" organization in the partnership, and with the strong support of the Council's leader, Jamil Zainaldin, became the primary public advocate for the encyclopedia. After the NGE's launch and its quick success, the Council came to view the NGE as perhaps the most prominent or high-profile of its programs, and continued to support the NGE through increased fundraising and efforts. After the initial funds were exhausted, the Council began supporting NGE directly out of its own operating budget (for one and a half FTEs). Additionally, by late 2009 the Council was providing office space for the NGE staff, now down to two employees.

At the start, GALILEO agreed to provide and maintain a server for the NGE. As NGE editors became more experienced and proficient with the publishing tool (text editor), they accumulated numerous requests for modifications to improve its functionality and efficiency. About a year and a half into working with the new site (and before its public launch), following fairly extensive and difficult modifications of the database, a decision was made to port the NGE code base from eXcelon to a better-performing database and to make extensive use of Java and related technologies for both presentation graphics as well as for such background functionality as statistics aggregation and processing. Oracle was selected because it was the standard set by the USG Office of Information Technology, and because USG already had a contract for the Oracle database. (The University System of Georgia covered the cost of the Oracle database.) The major task of porting the old database to a new, fully relational one was handled by GALILEO. After the NGE decided not to renew its maintenance agreement with the site developer, GALILEO by default took on a more substantial support role (with GALILEO recognizing the high-profile nature of the project). Soon GALILEO began covering the costs for a programmer full time, eventually reducing this support to half-time and then expecting to reduce support to one-quarter time. In reality, support requires about 80 % of a programmer's time. GALILEO still provides ongoing system administration support, equipment maintenance, and help-desk support, in addition to a programmer's time. Because of the NGE's aging technology, maintaining the site has been increasingly difficult and demanding, and GALILEO's in-kind contributions have dramatically exceeded the expectations for this component of the project.

After the NGE reached its goals for the first phase of content development, a reduction of the staff became necessary, as certain tasks and responsibilities became unnecessary and funding decreased. Since 2006 the NGE has employed only two FTEs, with periodic part-time assistance, who sustain and update the site's content and oversee or coordinate special grant-funded improvements to the site's content and functionality.

Benefits/Challenges

The benefits of spreading out among partners the risks of sustaining such an expensive and long-term project have well surpassed any negative aspects to partnership. No single partner could have provided the total resources necessary for the project to thrive. All partners have been able to "claim" the NGE as a successful, high-profile public project.

At times, decision-making can be a slow process, with implementation even slower. Occasionally, the organizations may have different priorities. Partners with particular expertise or greater resource investment in a certain area tend to have greater influence over decisions made in that area, and the result can be that what is tried-and-true is preferred over what is new or experimental. Another challenge for the NGE may include project

partner "fatigue," as the project enters its twelfth year and the resources and priorities of the partners invariably change or undergo reevaluation, particularly as personnel in the partner organizations come and go.

2. Content

Initial Development

The partners in the project believed that the New Georgia Encyclopedia should cover a wide spectrum of subjects rather than focus, as many state and regional encyclopedias have, on history and culture. As a humanities initiative, NGE should provide a comprehensive survey of history, literature, and the arts, but also reach beyond those disciplines to provide in-depth information on such areas as agriculture, commerce, education, geography, the environment, government, and science.

Those planning the encyclopedia also believed that NGE should examine issues affecting the state's present and future as well as its past. Because Georgia is one of the fastest-growing states in the nation, documenting its story in any full sense means including articles on the problems and opportunities created by this rapid growth. The partners agreed, then, that the encyclopedia should address such issues as urban sprawl; air and water pollution; the economic and social impact of new immigrant populations; political shifts and governmental policy; new trends in architecture, music, and visual arts; and the state's expanding role in TV/movie production and broadcast journalism. It was also extremely important that the content be accessible to students and general readers as well as to those with more specialized knowledge.

Because citizen input was vital, a number of "town meetings" were held around the state. The editors invited active, informed individuals representing all aspects of the community—governmental, cultural, business, philanthropic, and education—to attend the meetings and to share their thoughts about topics that should be covered in the encyclopedia.

In the early months of the project, the editorial staff evaluated other encyclopedias, consulted with editors and publishers of other large reference projects, and traveled within the state to promote the project and to build partnerships that would support the encyclopedia in significant and essential ways. In meetings with college and university faculty and leaders throughout the state, the editors identified potential contributors and documentary materials that could be incorporated into or linked to the online encyclopedia. Libraries, museums, research centers, institutes, and other cultural organizations in the state were invited to become Institutional Partners, and as such they committed to share the riches of their archival collections and the expertise of their curators and archivists with the project. Institutional Partners agreed to allow NGE to use multimedia materials for free or for a greatly reduced fee.

Individuals with expertise in particular subject areas were invited to serve as editorial advisors whose role would be to act as section editors and assist the NGE staff in identifying, soliciting, and evaluating articles for the encyclopedia. The Georgia Department of Economic Development agreed to provide assistance in developing a section on tourism and recreation. Together these individuals served as the project's editorial committee. After the completion of the first phase of content, most of the section editors resigned their posts. NGE staff editors now handle the initial review of new articles but often rely on project partners, NGE authors, former section editors, college professors, and others with expertise to help evaluate an entry's suitability.

Editorial Management

Once a completed article is received, it undergoes a lengthy editorial process. At the same time, the author and/or an NGE editor researches potential multimedia objects to accompany

the text. Following editorial review, copyediting, fact checking, final editing, and a last staff review of article, links, media, and captions, the entry is published to the live site.

Partnership with the University of Georgia Press and the University System of Georgia has allowed the NGE to secure university reference librarians as fact checkers. One librarian served as fact-checking coordinator and organized a pool of checkers, along with encyclopedia entries and deadlines. To verify the factual content of entries, checkers consult multiple sources, including primary sources when practical. At the end of the first phase of content development, the tasks handled by the fact-checking coordinator were absorbed by the NGE staff.

Entries are commissioned as "work-for-hire," and contracts, based on a Press contract template, are executed by NGE staff. The Georgia Humanities Council and the Press jointly hold copyright to all text on the site, and copyright is registered for the site as a whole. Copyright for multimedia objects is retained by the donor institution, as are the rights for the reproduction of multimedia.

Other Editorial Concerns

External linking. For external linking to other websites, editors have given preference to reliable, nonprofit sites and have mostly avoided personal pages and commercial sites.

Appropriateness. In deciding whether a topic is appropriate for the NGE, the topic's relevance to Georgia is considered, which is not always easy to determine. Editors have intentionally excluded entries on politicians currently holding office; topics that are too "newsy" and would require the passing of time in order to benefit from a historical perspective; entries that would disrupt the balance of coverage on the site (unless additional articles necessary for balance could also be written).

Change management. Any user may contact the NGE, through the site's contact page, with notice of a correction. Typos are corrected as soon as possible. Reported errors of fact or interpretation are handled on a case-by-case basis. Typically, the fact-checking materials for an entry are consulted first for more information. An editor may consult additional sources to verify the need for a correction, and if the matter may be simply resolved, she makes any necessary changes. Some error investigations require consulting with the original author and/or the academic editor of the NGE (as the years go by, involvement with the original author can become increasingly difficult). Often, the content in question must be returned to a fact checker for verification. Frequently, a correction made to one article must also be made elsewhere in the encyclopedia, so great care must be taken to avoid allowing conflicting information to exist within the site, which can undermine the site's authority.

Many entries require occasional updates, and some will require revision as new information emerges. Often, after an entry has been online for a number of years, NGE editors may ask individual authors or institutions whether their entry requires updating. The NGE uses a reporting feature of its publishing tool to send editors monthly email reminders about upcoming revisions (e.g., an institutional name change to take place on a certain date; the 50th anniversary of an approaching civil rights event). Frequent additions are necessary to entries on living people. NGE editors also keep track of the need to add new suggested readings to entries, new or changed URLs in the external links, and new media. Changes that require an updating to metadata are reported monthly to the Digital Library of Georgia, who maintains the NGE's metadata in a separate database.

From its inception, the NGE was intended to be a "living" project, in that the entries would be maintained and updated or added to as warranted over time. Editors are presently in the process of planning section updates, whereby each section or subsection would be reviewed as a whole for needed updates, new readings, new links among content, and so forth.

Finally, each entry carries the date on which it was published, or the date on which it was last updated.

Metadata. The NGE did not initially include metadata with its content. Through a grant-funded partnership with the Digital Library of Georgia (DLG), which is a project of GALILEO and the University of Georgia Libraries, a metadata librarian was hired to create metadata records for NGE's multimedia objects and articles using the Dublin Core metadata standard. DLG holds the metadata in its own database, from which the NGE search engine can access the metadata. DLG thus is able to maintain NGE's metadata records and make the information available to DLG's own users, who are able to search across multiple collections, including the encyclopedia, on the DLG site.

Post-Phase One Development

Since the initial phase of content development was completed, the NGE has periodically taken advantage of opportunities to partner with others to expand content coverage in certain subject areas. A partnership with the Digital Library of Georgia as part of the Civil Rights Digital Library initiative resulted in more than 30 new articles, as well as images and video footage, that enhanced the NGE's coverage of the civil rights movement. Working with a professor in the University of Georgia's College of Public Health resulted in several new entries, written by students, on biotechnology and environmental management. Similarly, NGE staff worked with two University of Georgia history professors to cover more deeply the Civil War in the state, in advance of the approaching Civil War sesquicentennial.

Repurposing Content

In 2007 *The New Georgia Encyclopedia Companion to Georgia Literature* was published by the University of Georgia Press and Georgia Humanities Council. The book's content was drawn entirely from the NGE website. In fall 2011 *The Civil War in Georgia: A New Georgia Encyclopedia Companion* will be co-published by the Press and the Council. This book will contain the essential story of the war in the state, again with content drawn from the NGE.

3. Sustainability

The concept of sustainability is one that the NGE staff and its project partners did not fully grasp when the encyclopedia was first created. Originally, the project had a plan to publish "phase one" of the content, then to "maintain" that content for the foreseeable future. As the staff neared its goal of completing phase one, however, the plan changed to develop a phase two of content, and NGE editors began compiling lists of potential topics.

In 2006, two years after NGE's launch and some five years into the content development, virtually all of the NGE section editors had moved on from the encyclopedia project. Most of the project's grant money was depleted. The project no longer had a budget for paying authors a heretofore nominal fee for contributions.

At this point, the project phased out two of its four staff positions. The positions lost were those heavily involved with the "front-end" NGE business--including working with section editors on their sections; preparing contracts for authors; requesting author payments; processing new entries received from authors; sending new entries out for review to section editors and the general editor; and general correspondence. These functions were absorbed as necessary by remaining staff. Additionally, the project's part-time programmer (paid for by GALILEO) left to take a job elsewhere.

Phase two of encyclopedia content was tabled to give the project more time to consult with outsiders on topics, to take care of the more pressing deadlines associated with a variety of other efforts in which NGE was engaged, and due to the shrinking time available to the smaller staff to devote to content development while still maintaining the existing body of content.

Two major aspects of keeping the NGE up and running began to consume increasingly more energy and resources: (1) as the software running our site aged, maintenance became more

challenging, and (2) as more and more content (articles and multimedia) was added to the site, greater effort was required to keep that content up-to-date and relevant.

Eventually, the project partners came to understand that the technology running the NGE site needed to be refreshed. In 2010 the database software running NGE (Oracle 9i) became officially obsolete when Oracle announced it would cease providing security updates for that version. GALILEO programmers have been upgrading the NGE to a newer version of Oracle, but the site has now reached its limits until a new platform can be acquired.

The NGE and its project partners at the outset imagined a site that could be easily "maintained" in perpetuity. No one anticipated how such a website could evolve to become much more than a place users visited to obtain information or meaning--that a website could become a dynamic, not passive, place for interactivity. Gradually, the NGE editors came to understand that the users of the website must be our *partners*; that if we fail to determine what their needs are and how they are using and will use the authoritative content that the NGE provides, then those users will simply turn elsewhere for information.

Moving forward to sustain the NGE requires making decisions about technology, content, features, and functionalities that keep the project in a stable but flexible position. Allowing for different ways of accessing NGE content, engaging and being responsive to users, and collaborating with our largest constituencies to fill in the gaps in our state's history, culture, and communities are critical to the NGE's plan for sustaining itself in the digital environment.

The NGE staff is now in the early planning stages for a "Next Generation NGE," which will involve migrating the NGE content to a new platform and positioning the project to be sustainable over time, especially with regard to technology. Research, user surveys, focus groups, and consulting with colleagues have yielded a few viable options. The next steps are to investigate the options further to find the solution that best fits, and to secure funding to carry out the implementation of the Next Generation.

—Kelly P. Caudle